Etude de deux problèmes de contrôle stochastique : Put Américain avec dividendes discrets et principe de programmation dynamique avec contraintes en probabilités

Maxence Jeunesse

THÈSE

présentée pour l'obtention du titre de

Docteur de l'Université Paris-Est

Spécialité : Mathématiques Appliquées

par **Maxence Jeunesse**

Sujet : Etude de deux problèmes de contrôle stochastique :
Put Américain avec dividendes discrets
et principe de programmation dynamique avec contraintes en probabilités

Soutenue le 29 01 2013
devant le jury composé de :

Rapporteurs :	Bruno Bouchard
	Stéphane Villeneuve
Examinateurs :	Pierre Carpentier
	Pierre Henry-Labordère
	Damien Lamberton
Directeurs de thèse :	Jean-Philippe Chancelier
	Benjamin Jourdain

Ces travaux ont bénéficié du support de la chaire "Risques Financiers", Fondation du Risque.

Au terme de ces trois années de thèse, je souhaite tout d'abord remercier mes deux directeurs de thèse Benjamin Jourdain et Jean-Philippe Chancelier pour leur encadrement sans faille et pour leurs précieux conseils toujours dispensés sans avarice.

Je tiens à remercier Bruno Bouchard et Stéphane Villeneuve d'être mes rapporteurs. Mes remerciements s'adressent également à Pierre Carpentier, Pierre-Henry Labordère et Damien Lamberton pour avoir accepté de faire partie de mon jury de thèse.

Je voudrais mentionner combien je suis redevable à tous les professeurs que j ?ai eu d'avoir influencé mes goûts mathématiques et plus particulièrement Nicole El Karoui et Sylvie Méléard.

J'aimerais exprimer toute ma gratitude à la chaire "Risques financiers", Fondation du Risque pour son support financier.

Ma reconnaissance s'adresse aussi à Jean-François Delmas pour m ?avoir permis de réaliser ma thèse dans d ?excellentes conditions au sein du laboratoire CERMICS de l ?Ecole des Ponts Paristech.

Merci à Nathalie Quelleu et Catherine Baccaert pour leur sens de l ?organisation. Merci à tous les chercheurs du laboratoire où j ?ai été si bien accueilli, en particulier Aurélien Alfonsi et Bernard Lapeyre qui ont livré de pertinentes analyses de problèmes économiques.

J'adresse également mes remerciements à José, mon co-bureau durant ces trois ans, ainsi qu'à Olivier, Patrick, Julien, Abdel, Stefano, Dylan, Fabien, Rémi qui ont su écouter mes nombreuses interrogations mathématiques, et qui ont contribué à créer une atmosphère chaleureuse lors des pauses cafés postprandiales avec Christelle, Adela, Charline, Pierre, Laurent et tous les autres doctorants et post-doctorants.

Fabrice, Julien, Vincent, Kévin, Yoann et tous les autres, merci d ?avoir manifesté tant d ?intérêt à l ?aboutissement de ce projet.

Merci à toute ma famille pour leur bienveillance, Josiane, Jean-Pierre, Gilbert, Nelly, Alexandra, et ceux qui sauront se reconnaître.

J'ai une pensée particulière pour ma grand-mère.

Enfin, je remercie mes parents, mon frère Paulien, et Mariette pour leur soutien indéfectible.

Sur un problème d'arrêt optimal et un résultat de programmation dynamique

Résumé : Dans cette thèse, nous traitons deux problèmes de contrôle optimal stochastique. Chaque problème correspond à une Partie de ce document. Le premier problème traité est très précis, il s'agit de la valorisation des contrats optionnels de vente de type Américain (dit Put Américain) en présence de dividendes discrets (Partie I). Le deuxième est plus général, puisqu'il s'agit dans un cadre discret en temps de prouver l'existence d'un principe de programmation dynamique sous des contraintes en probabilités (Partie II). Bien que les deux problèmes soient assez distincts, le principe de programmation dynamique est au coeur de ces deux problèmes.

La relation entre la valorisation d'un Put Américain et un problème de frontière libre a été prouvée par McKean. La frontière de ce problème a une signification économique claire puisqu'elle correspond à tout instant à la borne supérieure de l'ensemble des prix d'actifs pour lesquels il est préférable d'exercer tout de suite son droit de vente. La forme de cette frontière en présence de dividendes discrets n'avait pas été résolue à notre connaissance. Sous l'hypothèse que le dividende est une fonction déterministe du prix de l'actif à l'instant précédant son versement, nous étudions donc comment la frontière est modifiée. Au voisinage des dates de dividende, et dans le modèle du Chapitre 3, nous savons qualifier la monotonie de la frontière, et dans certains cas quantifier son comportement local. Dans le Chapitre 3, nous montrons que la propriété du smooth-fit est satisfaite à toute date sauf celles de versement des dividendes. Dans les deux Chapitres 3 et 4, nous donnons des conditions pour garantir la continuité de cette frontière en dehors des dates de dividende.

La Partie II est originellement motivée par la gestion optimale de la production d'une centrale hydro-electrique avec une contrainte en probabilité sur le niveau d'eau du barrage à certaines dates. En utilisant les travaux de Balder sur la relaxation de Young des problèmes de commande optimale, nous nous intéressons plus spécifiquement à leur résolution par programmation dynamique. Dans le Chapitre 5, nous étendons au cadre des mesures de Young des résultats dûs à Evstigneev. Nous établissons alors qu'il est possible de résoudre par programmation dynamique certains problèmes avec des contraintes en espérances conditionnelles. Grâce aux travaux de Bouchard, Elie, Soner et Touzi sur les problèmes de cible stochastique avec perte contrôlée, nous montrons dans le Chapitre 6 qu'un problème avec contrainte en espérance peut se ramener à un problème avec des contraintes en espérances conditionnelles. Comme cas particulier, nous prouvons ainsi que le problème initial de la gestion du barrage peut se résoudre par programmation dynamique.

Mots-clés : Arrêt optimal, Options Américaines, Dividendes, Frontière d'exercice, Propriété du smooth-fit, Processus de Lévy, Contrôle optimal stochastique, Consistence dynamique, Programmation dynamique.

Study of an optimal stopping problem and a result of dynamic programming principle

Abstract : In this thesis, we address two problems of stochastic optimal control. Each problem constitutes a different Part in this document. The first problem addressed is very precise, it is the valuation of American contingent claims and more specifically the American Put in the presence of discrete dividends (Part I). The second one is more general, since it is the proof of the existence of a dynamic programming principle under expectation constraints in a discrete time framework (Part II). Although the two problems are quite distinct, the dynamic programming principle is at the heart of these two problems.

The relationship between the value of an American Put and a free boundary problem has been proved by McKean. The boundary of this problem has a clear economic meaning since it corresponds at all times to the upper limit of the asset price above which the holder of such an option would exercise immediately his right to sell. The shape of the boundary in the presence of discrete dividends has not been solved to the best of our knowledge. Under the assumption that the dividend is a deterministic function of asset prices at the date just before the dividend payment, we investigate how the boundary is modified. In the neighborhood of dividend dates and in the model of Chapter 3, we know what the monotonicity of the border is, and we quantify its local behavior. In Chapter 3, we show that the smooth-fit property is satisfied at any date except for those of the payment of dividends. In both Chapters 3 and 4, we are able to give conditions to guarantee the continuity of the border outside dates of dividend.

Part II was originally motivated by the optimal management of the production of an hydro-electric power plant with a probability constraint on the reservoir level on certain dates. Using Balder's works on Young's relaxation of optimal control problems, we focus more specifically on their resolution by dynamic programming. In Chapter 5, we extend results of Evstigneev to the framework of Young measures. We show that dynamic programming can be used to solve some problems with conditional expectations constraints. Through the ideas of Bouchard, Elie, Soner and Touzi on stochastic target problems with controlled loss, we show in Chapter 6 that a problem with expectation constraints can be reduced to a problem with conditional expectation constraints. Finally, as a special case, we show that the initial problem of dam management can be solved by dynamic programming.

Keywords : Optimal stopping, American options, Dividends, Early exercise boundary, Smooth contact property, Lévy processes, Stochastic optimal control, Dynamic consistency, Dynamic programming.

Contents

Avant-propos

Cette thèse porte sur l'étude de deux problèmes de contrôle stochastique.

Un problème de contrôle stochastique ou problème d'optimisation stochastique est un problème générique où un décideur souhaite minimiser un coût (respectivement un joueur souhaite maximiser un gain) futur qui est par nature incertain, d'où le mot *stochastique*. Ses décisions modifient les chances de réalisation de la valeur du coût (respectivement du gain), et c'est en ce sens qu'il s'agit d'*optimisation*. Pour qualifier les actions du décideur (ou du joueur), on utilise indifféremment le mot décisions ou *contrôles*.

Ce document est organisé en deux parties. Les travaux menés sous la direction de Benjamin Jourdain sont exposés dans la Partie I. Ceux menés sous la direction de Jean-Philippe Chancelier correspondent à la Partie II. Bien que le principe de programmation dynamique énoncé par Bellman [Bel54] imprègne de manière plus ou moins évidente les deux parties de ce document, nous pensons qu'il est préférable de séparer l'introduction en deux chapitres distincts. Ainsi, le Chapitre 1 est une introduction à la Partie I, tandis que le Chapitre 2 est une introduction à la Partie II.

1

Problèmes d'arrêt optimal

L'objectif de ce Chapitre est de donner un bref résumé des résultats fondamentaux de la théorie de l'arrêt optimal ainsi qu'un aperçu de certains des travaux récents dans ce domaine. Dans un premier temps, nous introduirons le concept d'enveloppe de Snell dans un cadre séquentiel puis à temps continu. Ce concept est au centre de la théorie de l'arrêt optimal. Dans un second temps, nous présenterons le problème de la valorisation des contrats optionnels de vente de type Américain d'un actif financier. Enfin, nous résumerons les résultats que nous avons obtenu lorsque nous considérons qu'à certaines dates connues à l'avance, l'actif financier sous-jacent verse des dividendes discrets. Le Chapitre 3 consacré au cas où le sous-jacent évolue suivant le modèle de Black-Scholes en dehors des dates de dividendes a été publié dans *Stochastic Processes and their Applications* ([JJ12]).

1.1 Problèmes généraux d'arrêt optimal

Un problème d'arrêt optimal est un problème d'optimisation stochastique où un joueur essaie de maximiser son espérance de gain en décidant soit de continuer à jouer, soit de s'arrêter. Le premier problème de ce genre a été énoncé par Wald [Wal47] dans le cadre d'un test séquentiel en statistiques. Snell quant à lui a considéré le problème général d'arrêt d'un processus à temps discret [Sne52]. Il introduisit le concept d'enveloppe qui porte son nom, et ainsi caractérisa la solution en termes de martingales. Pour un traitement plus récent de ces sujets, et pour y trouver de nombreux exemples, nous nous référons à l'excellent livre de Peskir et Shiryaev [PS06].

1.1.1 Enveloppe de Snell

Temps discret

Soit $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\geq 0}, \mathbb{P})$ un espace de probabilité muni d'une filtration discrète, et soit $G = (G_n)_{n\geq 0}$ un processus adapté réel que nous supposerons borné pour des questions de simplicité dans les hypothèses des résultats suivants. Nous utiliserons la notation classique $a \wedge b = \min(a, b)$ pour tous réels a, b. L'*horizon* du problème est le dernier temps auquel nous pouvons encore jouer. Cet horizon peut être fini ou infini. Lorsque l'horizon est infini, il s'agit de trouver un temps d'arrêt $\tau^\star$ de la filtration $(\mathcal{F}_n)_{n\geq 0}$ tel que pour tout temps d'arrêt τ, $\mathbb{E}\left[G_{\tau^\star}\right] \geq \mathbb{E}\left[G_\tau\right]$. Lorsque l'horizon est fini, disons $N \geq 0$, il s'agit de trouver un temps d'arrêt $\tau^\star$ tel que pour tout temps d'arrêt τ, $\mathbb{E}\left[G_{\tau^\star\wedge N}\right] \geq \mathbb{E}\left[G_{\tau\wedge N}\right]$. On définit alors *l'enveloppe de Snell* de G comme la plus petite surmartingale qui domine G_n pour tout $n \geq 0$ si l'horizon est infini et seulement pour tout n entre 0 et N lorsque l'horizon du problème est N. Dans ce premier chapitre, on notera U^N l'enveloppe de Snell de G pour le problème d'horizon $N \geq 0$ qui se construit récursivement de la façon suivante (cf [LL08]).

Theorem 1.1.1 (Ch.I, Th.1.2 [PS06]) *Soit $N \geq 0$, et soit :*

$$U_n^N = \begin{cases} G_N & \text{si } n = N \\ \max(G_n, \mathbb{E}\left[U_{n+1}^N | \mathcal{F}_n\right]) & \text{si } n = 0 \dots N-1 \end{cases} . \tag{1.1}$$

Alors U^N est l'enveloppe de Snell de G pour le problème d'horizon $N \geq 0$ (ou plus simplement l'enveloppe de Snell d'horizon N de G).

Et le temps d'arrêt $\tau^\star = \inf\left\{n \in [\![0, N]\!] : U_n^N = G_n\right\}$ est optimal *dans le sens où :*

$$U_0^N = \sup_{\tau\in(\mathcal{F}_n)_{n\geq 0}} \mathbb{E}\left[G_{\tau\wedge N}\right] = \mathbb{E}\left[G_{\tau^\star\wedge N}\right], \tag{1.2}$$

où la notation $\tau \in (\mathcal{F}_n)_{n\geq 0}$ signifie que τ est un temps d'arrêt de la filtration $(\mathcal{F}_n)_{n\geq 0}$. De plus, le processus arrêté $(S_{n\wedge\tau^\star})_{n\geq 0}$ est une martingale.

Cette façon de résoudre le problème est parfois appelé la méthode martingale, puisqu'aucune structure Markovienne n'est nécessaire pour définir les objets.

Nous faisons maintenant l'hypothèse qu'il existe une chaîne de Markov $X = (X_n)_{n\geq 0}$ homogène à temps discret et à valeurs dans un espace localement compact à base dénombrable que l'on notera E. Le gain est une fonction Borélienne φ de E dans $\mathbb{R}$, c'est-à-dire que pour tout $n \geq 0$, $G_n = \varphi(X_n)$. Nous introduisons alors la famille de probabilités $(\mathbb{P}_x)_{x\in E}$ telle que l'état initial de X est x sous $\mathbb{P}_x$. L'opérateur de transition de X est noté P, c'est-à-dire que pour une fonction f qui est intégrable pour la loi de X_1 sous $\mathbb{P}_x$, on a :

$$Pf(x) = \mathbb{E}_x\left[f(X_1)\right]. \tag{1.3}$$

On suppose que pour tout $x \in E$ et $n \geq 0$, $\mathbb{E}_x\left[\sup_{0\leq k\leq n} |\varphi(X_k)|\right]$ est fini. Et on introduit pour $N \geq 0$:

$$V^N(x) = \sup_{\tau\in\mathcal{F}_{\leq N}^X} \mathbb{E}_x\left[\varphi(X_\tau)\right] \tag{1.4}$$

où $\mathcal{F}^X_{\leq N}$ est l'ensemble des temps d'arrêt de la filtration généré par X qui sont plus petits ou égaux à N. Un temps d'arrêt τ de $\mathcal{F}^X_{\leq N}$ est *optimal* pour le problème (1.4) lorsque pour tout $x \in E$, $V^N(x) = \mathbb{E}_x[\varphi(X_\tau)]$. On peut maintenant énoncer le résultat suivant :

Theorem 1.1.2 (Ch.I, Th.1.7 [PS06]) *On a $V^0 \equiv \varphi$, on a que pour $n \geq 1$, la fonction valeur V^n définie par l'Equation* (1.4) *avec n à la place de N satisfait l'équation de* Wald-Bellman *:*

$$\forall x \in E,\ V^n(x) = \max\left(\varphi(x), PV^{n-1}(x)\right). \tag{1.5}$$

Soit $N \geq 0$ et $x \in E$. Soit U^N l'enveloppe de Snell d'horizon $N \geq 0$ de G. Alors pour $0 \leq n \leq N$, U^N_n est $\mathbb{P}_x$-presque surement egal à $V^{N-n}(X_n)$.

Le temps d'arrêt $\tau^\star := \inf\left\{n \in [\![0, N]\!] : V^{N-n}(X_n) = \varphi(X_n)\right\}$ est optimal pour le problème (1.4). *De plus, pour tout autre temps d'arrêt τ optimal pour le problème* (1.4), *on a $\mathbb{P}_x(\tau \geq \tau^\star) = 1$ pour tout $x \in E$.*

Ce résultat apparaît dans les travaux de Bellman [Bel52, Bel57]. La preuve de ce résultat peut être déduite des résultats de Dvoretsky [DKW52]. Un cas particulier de ce résultat était déjà mentionné dans des travaux plus anciens de Wald et al. [WW49]. Les problèmes d'arrêt optimal dans un cadre Markovien et à temps discret ont été traités dans une grande généralité par Dynkin [Dyn63]. Des résultats généraux encore à temps discret, mais avec une structure probabiliste sous-jacente plus compliquée se trouvent dans [CRS71].

Temps continu

Soit $(\Omega, \mathcal{F}, (\mathcal{F}_s)_{s\geq 0}, \mathbb{P})$ un espace probabilisé muni d'une filtration à temps continu $(\mathcal{F}_s)_{s\geq 0}$ qui est continu à droite et telle que $\mathcal{F}_0$ contienne tous les ensembles $\mathbb{P}$-négligeables. Soit $(G_s)_{s\geq 0}$ un processus adapté réel et borné. Le problème est encore de trouver un temps d'arrêt $\tau^\star$ de $(\mathcal{F}_s)_{s\geq 0}$ tel que $\mathbb{E}[G_{\tau^\star}] = \sup_{\tau \in (\mathcal{F}_s)_{s\geq 0}} \mathbb{E}[G_\tau]$ où le supremum est pris sur tous les temps d'arrêt de la filtration $(\mathcal{F}_t)_{t\geq 0}$. Dans le but de définir un problème dynamique, on introduit le concept du problème partant de $t \geq 0$. Pour ce faire, nous aurons besoin de la notion d'*essentiel supremum* dont nous rappelons la définition dans un souci pédagogique. C'est le concept approprié pour obtenir des résultats qui seront similaires à ceux obtenus dans le cas où le temps était discret.

Définition 1.1 (Essentiel supremum). *Soit $(\Omega, \mathcal{G}, \mathbb{P})$ un espace de probabilité. Soit $\{Z_\alpha : \alpha \in I\}$ une famille de variables aléatoires réelles indexées par une ensemble arbitraire I. Alors il existe une variable aléatoire $Z^\star$ à valeurs dans $\mathbb{R} \cup \{+\infty\}$ qui est $\mathbb{P}$-presque sûrement unique et qui vérifie les trois propriétés suivantes :*

(a) il existe un ensemble dénombrable $J \subset I$ tel que $Z^\star = \sup_{\alpha \in J} Z_\alpha$,

(b) pour tout $\alpha \in I$, $\mathbb{P}(Z_\alpha \leq Z^\star) = 1$,

(c) pour tout autre variable aléatoire $\tilde{Z}$ vérifiant (b), on a $\mathbb{P}(Z^\star \leq \tilde{Z}) = 1$.

*La variable aléatoire $Z^\star$ est l'*essentiel supremum *de cette famille et on le note* $\operatorname{ess\,sup}_{\alpha \in I} Z_\alpha$.

Il s'agit donc pour le problème partant de $t \geq 0$ de trouver $\tau^\star \in (\mathcal{F}_s)_{s \geq 0}$ tel que $\mathbb{P}(\tau^\star \geq t) = 1$ et :

$$\mathbb{E}\left[G_{\tau^\star} | \mathcal{F}_t\right] = \operatorname*{ess\,sup}_{\tau \in (\mathcal{F}_s)_{s \geq 0} : \mathbb{P}(\tau \geq t) = 1} \mathbb{E}\left[G_\tau | \mathcal{F}_t\right]. \tag{1.6}$$

Grâce à ce concept, nous définissons *l'enveloppe de Snell U^T* (resp. $U^{+\infty}$) de G pour un horizon fini $T \geq 0$ (resp. infini) comme une version du processus $(\operatorname{ess\,sup}_{\tau \in [t,T]} \mathbb{E}\left[G_\tau | \mathcal{F}_t\right])_{t \in [0,T]}$ (resp. $(\operatorname{ess\,sup}_{\tau \geq t} \mathbb{E}\left[G_\tau | \mathcal{F}_t\right])_{t \geq 0}$) où pour tous $a, b \in [0, +\infty]$ la notation $\tau \in [a, b]$ est un abus pour désigner que le supremum est pris sur tous les temps d'arrêt τ de la filtration $(\mathcal{F}_t)_{t \geq 0}$ qui prennent leurs valeurs $\mathbb{P}$-presque sûrement dans $[a, b]$. Le résultat suivant est essentiel.

Theorem 1.1.3 (Ch.I, Th.2.2 [PS06]) *Soit $T \in [0, +\infty]$, en prenant la convention $G_{+\infty} = 0$. Supposons que $\mathbb{E}\left[\sup_{0 \leq t \leq T} |G_t|\right] < \infty$. Nous pouvons choisir une version continue à droite de U^T. Soit $t \geq 0$, on définit alors $\tau^\star = \inf\left\{s \in [t, T] : U^T_s = G_s\right\}$. Si $\mathbb{P}(\tau^\star < +\infty) < 1$ alors il n'y a pas de temps d'arrêt qui soit optimal pour le problème* (1.6) *partant de t. Si $\mathbb{P}(\tau^\star < +\infty) = 1$, alors $\tau^\star$ est optimal pour le problème* (1.6) *partant de t. De plus, pour tout autre temps d'arrêt τ optimal pour le problème* (1.6) *on a $\mathbb{P}(\tau^\star \leq \tau) = 1$. Enfin, le processus $(U^T_{s \wedge \tau^\star})_{s \in [t,T]}$ est une martingale.*

Un grand nombre de résultats plus précis sont établis dans les livres de Shiryaev [Shi08] et de El Karoui [EK81, Ch.II]. Dans l'article référent de El Karoui, Pardoux et Quenez [EKPQ97], une caractérisation de l'enveloppe de Snell est obtenue comme solution d'une équation différentielle stochastique rétrograde.

Supposons que T soit fini et que l'espace de probabilité soit l'espace de Wiener en dimension d. On note W le mouvement Brownien d-dimensionnel associé au processus canonique. Comme précédemment, $\tau \in [t, T]$ est un abus pour désigner les temps d'arrêt de la filtration générée par W et telle que $\mathbb{P}(t \leq \tau \leq T) = 1$. Soit ξ une variable aléatoire réelle $\mathcal{F}_T$-mesurable telle que $\mathbb{E}\left[|\xi|^2\right] < \infty$ et soit L un processus réel adapté à $(\mathcal{F}_t)_{0 \leq t \leq T}$ et tel que $\mathbb{E}\left[\sup_{0 \leq t \leq T} |L_t|^2\right] < \infty$. On énonce alors le résultat suivant :

Proposition 1.1.4 (Ch.6, Prop.6.5.3 [Pha09]) *Soit Y l'enveloppe de Snell du processus $t \in [0, T] \mapsto L_t \mathbf{1}_{\{t < T\}} + \xi \mathbf{1}_{\{t = T\}}$ i.e pour $t \in [0, T]$,*

$$Y_t = \operatorname*{ess\,sup}_{\tau \in [t,T]} \mathbb{E}\left[L_\tau \mathbf{1}_{\{\tau < T\}} + \xi \mathbf{1}_{\{\tau = T\}}\right].$$

Alors il existe un processus adapté à trajectoires croissantes K tel que $K_0 = 0$ et un processus adapté Z à valeurs dans $\mathbb{R}^d$ tel que $\mathbb{E}\left[\int_0^T |Z_t|^2\, dt\right] < \infty$ et on a $\mathbb{E}\left[\sup_{0\le t\le T} |Y_t|^2\right] < \infty$ et :

$$\begin{aligned} Y_t = \xi + K_T - K_t - \textstyle\int_t^T Z_s.dW_s \; 0 \le t \le T \\ Y_t \ge L_t,\, 0 \le t \le T \\ \textstyle\int_0^T (Y_t - L_t) dK_t = 0 \end{aligned} \tag{1.7}$$

Le triplet (Y, Z, K) est appelé une solution de l'équation différentielle stochastique rétrograde réfléchie (1.7)*. Les hypothèses de cet énoncé assurent l'unicité de la solution.*

Si on fait l'hypothèse supplémentaire que $\xi = g(W_T)$ et $L_t = h(t, W_t)$ pour des fonctions déterministes h et g, on peut prouver que la solution (Y, Z, K) est reliée à l'une des solutions régulières v de l'inéquation variationnelle $\max(\partial_t v + \Delta_x v, h - v) = 0$ on $[0, T) \times \mathbb{R}^d$ et $v(T, .) = g$ sur $\mathbb{R}^d$. Ces résultats sont établis dans un cadre plus général dans [EKPQ97]. En effet, comme on va le voir dans la sous-section suivante, en présence d'une structure Markovienne, des résultats plus précis sur l'enveloppe de Snell peuvent être obtenus.

1.1.2 Principes variationnels et problèmes de frontière libre

Soit E un espace localement compact et séparable. Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité complet et soit X un processus de Markov homogène à temps continu à valeurs dans E dont les trajectoires sont continues à droites et ont des limites à gauches. On note son générateur infinitésimal L. Soit $\mathbb{P}_x$ la probabilité sous laquelle X part de $x \in E$. On fait l'hypothèse que le processus X rassemble toute l'information nécessaire pour prendre une décision. Ainsi, la filtration naturelle est celle générée par les trajectoires de X. Soit $\mathcal{F}_0$ la tribu sur Ω engendrée par X_0 et les ensembles $\mathbb{P}$-négligeables. On définit alors pour tout $t \ge 0$, $\mathcal{F}_t$ comme la tribu engedrée par $\mathcal{F}_0$ et $(X_s)_{s\le t}$. Puisque X a des trajectoires continues à droites, $(\mathcal{F}_t)_{t\ge 0}$ est une filtration continue à droite. Soit $\varphi : E \mapsto \mathbb{R}$ une fonction Borèlienne. Soit $T \in [0, +\infty]$ un horizon, avec la convention que $\varphi(X_{+\infty}) = 0$. Faisons l'hypothèse que pour tout $x \in E$, $\mathbb{E}_x\left[\sup_{0\le t\le T} |\varphi(X_t)|\right] < +\infty$. On s'intéresse alors à déterminer la fonction valeur $u^T : E \mapsto \mathbb{R}$ et le temps d'arrêt optimal $\tau^\star$ de $(\mathcal{F}_t)_{t\ge 0}$ tel que :

$$u^T(x) = \sup_{\tau\in(\mathcal{F}_t)_{t\ge 0}} \mathbb{E}_x\left[\varphi(X_{\tau\wedge T})\right] = \mathbb{E}_x\left[\varphi(X_{\tau^\star\wedge T})\right]. \tag{1.8}$$

Nous énonçons une version très affaiblie des théorèmes généraux qui existent. Ce sera cependant suffisant pour illustrer la suite de nos propos.

Theorem 1.1.5 *On suppose que X est un processus de Feller, et que φ est continue. Si T est fini, et si u^T est dans le domaine de* L *et est dérivable par rapport à T alors u est une solution de l'inéquation variationnelle :*

$$0 = \max(\varphi(x) - u^T(x), \mathrm{L}u^T(x) - \partial_T u^T(x)), \quad u^0 \equiv \varphi \tag{1.9}$$

Si $T = +\infty$ *et si* u^T *est dans le domaine de* L*, alors* $0 = \max(\varphi(x) - u^T(x), \mathrm{L}u^T(x))$.

Eléments de preuve Si T est fini. En prenant $\tau \equiv 0$, on a que pour tout $x \in E$, $\varphi(x) \leq u^T(x)$. A cause de la continuité de φ et de la propriété de Feller de X, par [PS06, Eq.(2.2.80)] nous avons la régularité suffisante sur u^T pour appliquer les résultats de Baxter Chacon [Mey78]. On obtient alors que le principe de programmation dynamique est valide, c'est-à-dire que pour tout $t \in [0, T]$ et $x \in E$, on a $u^T(x) \geq \mathbb{E}_x\left[u^{T-t}(X_t)\right]$. Par conséquent, si u^T est dans le domaine de L et est dérivable par rapport à T alors $\mathrm{L}u^T(x) \leq \partial_T u^T(x)$. Enfin, nous obtenons qu'une conséquence directe du Théorème 1.1.3 est que pour $x \in E$, l'inégalité $u^T(x) > \varphi(x)$ implique que $\mathrm{L}u^T(x) - \partial_T u^T(x) = 0$. En effet, pour $x \in E$ tel que $u^T(x) > \varphi(x)$, l'hypothèse de continuité de φ garantit que $D = \left\{x \in E : u^T(x) > \varphi(x)\right\}$ est un ouvert. En prenant τ le temps de sortie d'une boule centrée autour de x et incluse dans D, on a, par le principe de programmation dynamique et puisque $(u^{T-t}(X_t))_{t\in[0,T]}$ est l'enveloppe de Snell d'horizon T de $(\varphi(X_t))_{t\in[0,T]}$, $u^T(x) = \mathbb{E}\left[u^{T-\tau}(X_\tau)\right]$ en utilisant le Théorème 1.1.3. Ceci prouve l'égalité précédemment annoncée.
Si $T = +\infty$, alors pour tout t fini, $u^{T-t} \equiv u^T$ et on réitère le raisonnement précédent.

Comme l'exemple suivant l'illustre, ce résultat n'est pas suffisant pour caractériser la fonction u^T définie par (1.8). En effet, dans de nombreux cas, on obtient les premières propriétés de régularité de u^T à partir de celles de φ sans trop de difficultés. Mais ces propriétés ne permettent pas de trouver une unique solution à l'inéquation variationnelle du précédent Théorème. Heureusement, lorsque le comportement de la fonction valeur u^T est connu sur la frontière de l'ensemble $\left\{x \in E : u^T(x) = g(x)\right\}$, alors on a bien souvent l'unicité d'une solution satisfaisant une telle contrainte. En effet, dans beaucoup de problèmes où le but est de trouver plus ou moins explicitement la fonction valeur d'un problème d'arrêt optimal, on peut remarquer qu'ou bien la continuité au passage de la frontière, ou bien la propriété du smooth-fit sont suffisantes pour conclure. (Pour un traitement complet de cette dichotomie, nous renvoyons à [PS06, Ch.IV.Sec.9]). La propriété du *smooth-fit* est en fait simplement la propriété que la fonction valeur est continue et différentiable en x au passage de la frontière.

Illustration du smooth-fit

Dans le but d'illustrer comment tous les précédents résultats s'articulent, nous prenons un exemple particulier que nous retrouverons dans §1.2.1. Les grandes lignes pour trouver une solution explicite à un problème d'arrêt optimal sont très souvent les mêmes. Il s'agit tout d'abord de trouver des estimées à priori sur la fonction valeur. Ces estimées vont ensuite nous être utiles pour trouver la forme de la région où la fonction valeur u^T est égale à la fonction de gain φ. Arrivés à cette étape, la fonction valeur peut n'être pas encore déterminée. Dans ce cas, il suffit souvent pour complètement caractériser la fonction valeur prouver que celle-ci admet la propriété du *smooth-fit*. La solution est alors l'unique fonction qui satisfait à toutes les contraintes précédentes.

Soit $T = +\infty$, $r \geq 0$ et $\sigma > 0$. Soit S, le processus normalisé de Black Scholes i.e $(S_t = e^{(r-\sigma^2/2)t+\sigma B_t})_{t\geq 0}$ où B est un brownien sous $\mathbb{P}$. On note A le processus déterministe de drift constant égal à 1. On considère la paire $X = (X^{(1)}, X^{(2)})$ à valeurs dans $\mathbb{R}\times\mathbb{R}_+$ tq pour $x = (t, y) \in \mathbb{R}\times\mathbb{R}_+$, la loi de X sous $\mathbb{P}_x$ est celle de $(t+A, yS)$. On choisit $\varphi(x) = \varphi(x_1, x_2) = e^{-rx_1}(1-x_2)^+$ où $a^+ = \max(a, 0)$. Et notre objectif est déterminer la fonction u tel que pour tout $(t, y) \in \mathbb{R}\times\mathbb{R}_+$ on a :

$$u(t,y) = \sup_{\tau\in(\mathcal{F}_s)_{s\geq 0}} \mathbb{E}_{(t,y)}\left[e^{-rX_\tau^{(1)}}(1-e^{X_\tau^{(2)}})^+\right] = e^{-rt} \sup_{\tau\in(\sigma(S_u:u\leq s))_{s\geq 0}} \mathbb{E}_{(0,1)}\left[e^{-r\tau}(1-yS_\tau)^+\right]$$

Afin de simplifier les notations, on notera $\mathbb{P}$ au lieu de $\mathbb{P}_{(0,1)}$.

Dans un premier temps,

comme $t \geq 0 \mapsto e^{-rt}(1-yS_t)^+$ est bornée pour tout $y \geq 0$, on vérifie sans peine que $\mathbb{E}\left[\sup_{t\geq 0} e^{-rt}(1-yS_t)^+\right]$ est bornée par 1 pour tout $y \geq 0$. Ensuite à cause du choix particulier de φ, nous pouvons affirmer que u est le produit de $x \mapsto e^{-rx_1}$ et de la fonction v définie ci-dessous qui ne dépend que de x_2 :

$$v : y \in \mathbb{R}_+ \mapsto \sup_{\tau\in(\sigma(S_u:u\leq s))_{s\geq 0}} \mathbb{E}\left[e^{-r\tau}(1-yS_\tau)^+\right]. \tag{1.10}$$

Pour $y \leq y'$, on a pour tout $\tau \in (\sigma(S_u : u \leq s))_{s\geq 0}$, $\mathbb{P}(e^{-r\tau}(1-yS_\tau)^+ \geq e^{-r\tau}(1-y'S_\tau)^+) = 1$, donc v est décroissante. En remarquant que $(e^{-rt}S_t)$ est une martingale, on montre de manière analogue que la fonction $y \mapsto y + v(y)$ est croissante. Nous avons donc établi à priori que v est décroissante, 1-Lispchitz et borné par 1.

Dans un deuxième temps,

nous allons maintenant donner la forme de l'ensemble $D := \{y \in \mathbb{R}_+ : v(y) = (1-y)^+\}$. Tout d'abord, nous remarquons que $0 \in D$. Donc D n'est pas vide, ensuite nous remarquons en prenant $\tau = 1$ (et puisque $\sigma > 0$), que $[1, +\infty) \notin D$. Nous définissons alors $\kappa = \sup\{y \geq 0 : v(y) = (1-y)^+\}$. Par les remarques précédentes, $\kappa \in [0, 1]$. En utilisant alors la croissance de $y \mapsto y + v(y)$, ce qui est équivalent à la croissance de $y \mapsto v(y) - (1-y)^+$ sur $[0, 1]$, nous établissons que $D = [0, \kappa]$.

On sait en utilisant le Théorème 1.1.5 que sur D^c, $v(y) = B_1 y + B_2 y^{-\frac{2r}{\sigma^2}}$ pour deux constantes réelles B_1, B_2. Mais les hypothèses de bornitude de v implique que $B_1 = 0$. Ainsi, comme v est continue, on voit que $v(y) = (1-y)\mathbf{1}_{\{y<\kappa\}} + (1-\kappa)\left(\frac{y}{\kappa}\right)^{-\frac{2r}{\sigma^2}}\mathbf{1}_{\{y\geq\kappa\}}$.

Application de la propriété du smooth-fit

Nous savons que $\kappa \in (0,1)$ puisque v est bornée et strictement positive. Il reste à prouver que $\kappa = \frac{2r}{2r+\sigma^2}$ est la valeur correcte. Nous avons deux possibilités de preuve. Ou bien, on le montre par un argument de type vérification, c'est-à-dire que l'on calcule explicitement v en utilisant le fait que $\tau^\star = \inf\{t \geq 0 : yS_t \in D\}$ est optimal pour le problème (1.10). Ou bien, on essaie de prouver la propriété du *smooth-fit*. On rappelle que celle-ci établit que la fonction valeur est $\mathcal{C}^1$ au passage de la frontière ∂D. Ici, en fait, nous sommes exactement dans le cas d'application de [Pha09, Prop.5.2.1] et par conséquent, nous pouvons établir que v est $\mathcal{C}^1$ en κ. Par conséquent, on a que κ est l'unique solution positive de $v'(\kappa) = -\frac{2r}{\sigma^2}\frac{v(\kappa)}{\kappa} = -1$, c'est-à-dire la valeur annoncée.

Considérations historiques

Le premier à souligner la propriété du smooth-fit fut Mikhalevich dans [Mik58]. Puis d'autres auteurs ont remarqué l'importance de cette propriété (par exemple [Dyn63, McK65]).

L'équation sur u^T du Théorème 1.1.5 est une inéquation variationnelle. Pour un traitement rigoureux de ce vaste sujet, nous renvoyons au livre de Friedman [Fri88].

1.2 Contrats optionnels de type Américain

Depuis le travail singulier de McKean [McK65], la littérature sur la valorisation d'un contrat optionnel de type Américain n'a cessé de croître.

Nous rappelons qu'un contrat optionnel de type Américain est caractérisé par un payoff et une maturité. Si T est la maturité d'un contrat de ce genre, et $(G_t)_{t\in[0,T]}$ est le payoff, alors le détenteur de ce contrat peut, à tout instant $t \in [0,T]$, décider d'exercer son droit et être payé G_t. Un grand nombre d'auteurs ont prouvé que dans un modèle uni-dimensionnel de Black Scholes, le vendeur d'une telle option, dans le cadre d'une option de vente, peut construire un portefeuille autofinancé jusqu'à un certain temps d'arrêt puis ensuite arbitrer l'acheteur si ce dernier n'exerce pas son droit de vente à cet instant précis. Une option de vente se définit pour une maturité T comme un contrat optionnel de payoff $G_t = (K - xe^{(r-\frac{\sigma^2}{2})t+\sigma B_t})^+$, où B est un mouvement brownien, r représente le taux d'intérêt sans risque et σ la volatilité de l'actif. K est une constante appelée le strike. Ce contrat a une valeur (ou prix), on la note P_t, et elle est donnée à tout instant $t \in [0,T]$ par :

$$P_t = \operatorname*{ess\,sup}_{\tau\in[t,T]} \mathbb{E}\left[e^{-r\tau}\left(K - xe^{\left(r-\frac{\sigma^2}{2}\right)t+\sigma B_\tau}\right)^+ \middle| \mathcal{F}_t\right] \tag{1.11}$$

où le supremum est pris sur tous les temps d'arrêt de la filtration $(\mathcal{F}_t)_{t\geq 0}$ du mouvement Brownien B qui sont $\mathbb{P}$-presque sûrement à valeurs dans $[t,T]$. Il est intéressant de noter que cette valeur est la valeur maximale qu'un acheteur est prêt à payer pour acheter une option de ce type, puisqu'il s'agit exactement de toutes les valeurs initiales de portefeuilles autofinancés qui répliquent un contrat où l'acheteur aurait spécifié au moment de l'achat sa politique d'exercice donnée par τ. Le contrat de prix P_t à l'instant t est une option Américaine, et est appelé Put Américain de strike K et de maturité T sur le sous-jacent $(xe^{(r-\frac{\sigma^2}{2})t+\sigma B_t})_{t\in[0,T]}$.

Bensoussan [Ben84] a prouvé que dans de nombreux modèles de diffusion la valeur d'une option Américaine est reliée à la solution d'une inéquation variationnelle. Pham [Pha97] a généralisé ce lien dans le cadre de processus de diffusions avec sauts.

1.2.1 Puts Américains standards

Nous allons énoncer les résultats dans le modèle exponentiel de Lévy. On rappelle tout d'abord que dans un modèle exponentiel de Lévy, le processus du prix de l'actif à l'instant t partant de x en 0 est égal à xe^{X_t} pour X un processus de Lévy unidimensionnel, et il existe un taux d'intérêt sans risque $r \geq 0$. Rentrons maintenant dans les détails. Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé complet et soit X un processus de Lévy réel d'exposant caractéristique ψ définie par $\mathbb{E}\left[e^{\lambda X_t}\right] = e^{t\psi(\lambda)}$ pour $\lambda \in \Lambda$ où $\Lambda = \{\lambda \in \mathbb{C} : \Re(\lambda) \in \mathrm{I}\}$ et où I est un intervalle contenant $[0,1]$. Par la formule de Lévy-Khintchine, on a :

$$\psi : \lambda \in \Lambda \mapsto \psi(\lambda) = \frac{\sigma^2}{2}\lambda^2 + \gamma\lambda + \int_{\mathbb{R}} \left(e^{\lambda y} - 1 - \lambda y \mathbf{1}_{\{|y|\leq 1\}}\right)\nu(dy) \tag{1.12}$$

où (σ, γ, ν) est le triplet caractéristique de X. En particulier, ν est une mesure (de Lévy) sur $\mathbb{R}$, c'est-à-dire que $\int_{\mathbb{R}}(1 \wedge y)^2\nu(dy) < \infty$ et $\nu(\{0\}) = 0$. On fait l'hypothèse que $\psi(1) = r \geq 0$. Et par analogie avec le cas Black Scholes, on s'intéresse au Put Américain standard de strike K et de maturité T (avec $T, K > 0$ et finis) :

$$v : (T, x) \in \mathbb{R}_+^2 \mapsto \sup_{\tau\in[0,T]} \mathbb{E}\left[e^{-r\tau}\left(K - xe^{X_\tau}\right)^+\right] \tag{1.13}$$

où le supremum est pris sur tous les temps d'arrêt de la filtration générée par X qui sont $\mathbb{P}$-presque sûrement à valeurs dans $[0,T]$. On définit également le Put Américain standard perpétuel de strike K :

$$v : x \in \mathbb{R}_+ \mapsto \sup_{\tau\geq 0} \mathbb{E}\left[e^{-r\tau}\left(K - xe^{X_\tau}\right)^+\right] \tag{1.14}$$

où le supremum est pris sur tous les temps d'arrêt de la filtration générée par X.

Maturité infinie

Dans le cas où la maturité est infinie, on qualifie l'option Américain de perpétuelle. Nous reprenons ici les résultats de Mordecki [Mor02]. Afin de simplifier leur énoncé, nous supposerons $r > 0$. On définit alors $\mathfrak{e}_r$ une variable aléatoire indépendante de X, de loi exponentielle de paramètre r.

Theorem 1.2.1 (Th.2 [Mor02] Perpetual put options) *En posant* $I = \inf_{0 \leq s \leq \mathfrak{e}_r} X_s$ *et* $\kappa = K\mathbb{E}\left[e^I\right]$*, on a :*

$$v(x) = \begin{cases} K - x & si\ x \leq \kappa \\ \mathbb{E}\left[\left(K - x\dfrac{e^I}{\mathbb{E}\left[e^I\right]}\right)^+\right] & si\ x > \kappa \end{cases} \tag{1.15}$$

et le temps d'arrêt, $\tau^\star = \inf\left\{t \geq 0 : xe^{X_t} \leq \kappa\right\}$ *est optimal pour le problème* $v(x)$*.*

Dans le cas du modèle de Black Scholes et lorsque la volatilité de l'actif est $\sigma > 0$, nous avons vu que la valeur est donnée par :

$$v(x) = \begin{cases} K - x & \text{si } x \leq K\frac{2r}{2r+\sigma^2} \\ K\frac{\sigma^2}{2r+\sigma^2}\left(\frac{x}{K\frac{2r}{2r+\sigma^2}}\right)^{-\frac{2r}{\sigma^2}} & \text{si } x > K\frac{2r}{2r+\sigma^2} \end{cases} \tag{1.16}$$

En fait, Mordecki [Mor02, Th.4] a établi des formules analogues pour un processus de Lévy dont la mesure de saut a une forme particulière. Plus récemment, Mordecki et Salminen [MS07] ont donné des solutions explicites pour des fonctions plus générales que $x \mapsto (K - x)^+$ et pour une large classe de processus de Markov.

Maturité finie

Dans le cas du modèle de Black Scholes, le problème remonte à McKean [McK65] et Van Moerbeke [VM76] mais une liste presque complète des résultats sur ce problème se trouve dans l'article de Myneni [Myn92]. Pour le cas du Put Américain standard dans un modèle exponential de Lévy, beaucoup de résultats se trouvent dans les articles de Lamberton et Mikou [LM08, LM12b]. Nous allons rappeler ici les principaux résultats.

Proposition 1.2.2 (Prop.3.2 [LM08] Propriétés de régularité) *Pour tout* $T \geq 0$*, la fonction* $v(T, .)$ *est décroissante, convexe et* 1 *Lipschitz. Pour tout* $x \geq 0$*, la fonction* $v(., x)$ *est croissante. La fonction* v *est globalement continue sur* $\mathbb{R}_+^2$*.*

On fait l'hypothèse maintenant qu'au moins une des conditions suivantes est vérifiée.

a) $\sigma > 0$,

b) $\int_{(0,1]} |x|\, \nu(dx) = +\infty$,

c) $\nu(\mathbb{R}_-^\star) > 0$.

Et nous pouvons énoncer le résultat suivant.

Proposition 1.2.3 (Prop.4.1,Th.4.2 [LM08] Frontière d'exercice) *Il existe une fonction continue c de $\mathbb{R}_+^\star$ dans $(0,K]$ telle que pour tout $t>0$, $c(t)=\sup\{x\geq 0 : v(t,x)=(K-x)^+\}$. De plus pour tout $(T,x)\in\mathbb{R}^2$, le temps d'arrêt $\tau^\star = \inf\left\{t\geq 0 : xe^{X_t}\leq c(T-t)\right\}\wedge T$ est optimal pour $v(T,x)$. C'est-à-dire que $\left(e^{-rt\wedge\tau^\star}v(T-t\wedge\tau^\star, xe^{X_{t\wedge\tau^\star}})\right)_{t\geq 0}$ est une martingale pour $\mathbb{P}$ et $v(T,x)=\mathbb{E}\left[e^{-r\tau^\star}(K-xe^{X_{\tau^\star}})^+\right]$.*

La fonction c est dénommée la *frontière d'exercice*. Cela vient du fait que le détenteur d'un contrat de ce type, désirant exercer optimalement, doit exercer son option au premier instant où le prix de l'actif passe sous cette frontière.

Proposition 1.2.4 (Th.4.4 [LM08]) *Sous l'hypothèse supplémentaire que $\int(e^y-1)^+\,\nu(dy)\leq r$, la fonction c définie précedemment tend vers K lorsque t tend vers 0.*

Dans le cadre spécifique du modèle de Black-Scholes, Ekstrom ainsi que Chen et al. [Eks04, CCJZ08] ont prouvé que la frontière est convexe. Une preuve rigoureuse du caractère infiniment dérivable de la frontière sur $\mathbb{R}_+^\star$ a été établie par Chen et Chadam dans [CC06]. De plus, des estimées du comportement de la frontière au voisinage de la maturité ont été obtenus par Lamberton [Lam95] puis par Lamberton et Villeneuve [LV03].

Proposition 1.2.5 (Th.3 [Lam95]) *Avec les notations précédentes, on a :*

$$\lim_{\theta\to 0}\frac{K-c(\theta)}{\sigma K\sqrt{\theta\,|\ln\theta|}}=1. \tag{1.17}$$

Des travaux plus récents de Lamberton et Mikou [LM12a] ont montré une grande variété de comportements de la frontière au voisinage de la maturité dans le cadre des modèles exponentiels de Lévy.

Nous allons énoncer maintenant la propriété du smooth-fit.

On fait l'hypothèse qu'une des conditions suivantes est satisfaite :

1. X est à variation finie et $\int_{\mathbb{R}}(1-e^y)\nu(dy)<0$.
2. X est à variation finie et $\int_{\mathbb{R}}(1-e^y)\nu(dy)=0$ et $\int_{(-1,0)}\frac{|x|}{\int_0^{|x|}\nu([y,+\infty))dy}\nu(dx)=+\infty$.
3. X est à variation infinie.

Alors on a :

Proposition 1.2.6 (Th.4.1 [LM12b] Smooth-fit) *Pour tout $T>0$, la fonction $x\mapsto v(T,x)$ est continument dérivable sur $\mathbb{R}_+$.*

En fait, la Proposition précédente s'énonce de la façon suivante dans le langage de la théorie du potentiel. Si le point 0 est régulier pour $(-\infty, 0)$ (i.e $\mathbb{P}(\inf\{t > 0 : X_t < 0\} = 0) = 1$) alors la propriété du smooth-fit est satisfaite pour $v(T, .)$ à tout instant $T > 0$.

1.2.2 Résumé des résultats de la Partie I

Il est clair que la modelisation d'un actif financier en négligeant le versement des dividendes n'est pas réaliste. Nous avons donc été amené à considérer le cas stylisé décrit par Goettsche Vellekoop et Nieuwenhuis [GV11, VN11]. Nous nous intéressons donc à des options Américaines sur un actif qui verse des dividendes à des instants connus à l'avance et ce avant la maturité de l'option. Cette modélisation est pertinente dans le sens où les dates de paiement de dividendes sont données par les entreprises. Nous faisons l'hypothèse que le processus de prix entre deux dates de versement des dividendes a une évolution connue (de type exponentielle de Lévy ou plus spécifiquement Black-Scholes). On suppose également que le montant du dividende est une fonction déterministe du prix juste avant la date de versement de dividende. On suppose que cette fonction est positive, croissante et est 1-Lipschitz. Nos résultats étendent les résultats de Jourdain et Vellekoop [JV11]. En effet, sous certaines hypothèses assez générales, nous prouvons que la frontière est continue sur les intervalles de temps entre dates de versement de dividendes. Nous établissons des résultats sur la propriété du smooth-fit. Enfin, nous donnons des conditions sur la fonction de dividende pour garantir la monotonie locale au voisinage d'une date de versement des dividendes.

Construction récursive

Nous considérons le Put Américain de maturité T et de strike K sur un sous jacent S. Entre 0 et T, nous supposons qu'il y a I dates de versement de dividendes. C'est-à-dire que ce sous jacent verse un dividende à des instants déterministes $0 \leq t_d^I < t_d^{I-1} < \cdots < t_d^i < \cdots < t_d^1 < T$. A chaque date t_d^i, $S_{t_d^i} = S_{t_d^i-} - D_i(S_{t_d^i-})$, où $D_i(S_{t_d^i-})$ est la valeur du dividende payé. Comme dit précedemment, $D_i : \mathbb{R}_+ \to \mathbb{R}_+$ est une fonction croissante, positive nulle en zéro et 1-Lipschitz. On a donc que $x \mapsto x - D_i(x)$ est aussi croissante et positive.

On fait l'hypothèse que la valeur d'un tel contrat lorsque la valeur initiale de l'actif est x est donnée par :

$$P(x) = \sup_{\tau \in [0,T]} \mathbb{E}\left[e^{-r\tau}(K - S_\tau)^+\right] \tag{1.18}$$

où le supremum est pris sur tous les temps d'arrêt de la filtration de S et où l'on considère que $\mathcal{F}_0$ contient l'information sur les dates de versement des dividendes et $\mathbb{P}$ est une probabilité de pricing (la même que dans la sous-section précédente).

Nous faisons l'hypothèse qu'entre les instants de versement des dividendes, l'actif suit un modèle exponentiel de Lévy comme dans la sous-section précédente.

Nous allons poser $(\theta_d^i = t_d^i - t_d^{i+1})_{0\leq i\leq I-1}$ avec la convention $t_d^0 = T$. Puis nous allons récursivement définir des fonctions qui seront des intermédiaires pour caractériser la valeur $P(x)$.

Avec les notations de la sous-section précédente, nous posons $u_0 = v$. Et nous posons $c_0 = c$ pour la frontière d'exercice. Ensuite pour $i \in \{1, \dots, I\}$, on définit :

$$u_i : (\theta, x) \in \mathbb{R}_+ \times \mathbb{R}_+ \mapsto \sup_{\tau\in[0,\theta]} \mathbb{E}\left[\begin{array}{l} e^{-r\tau}(K - xe^{X_\tau})^+\mathbf{1}_{\{\tau<\theta\}} \\ + \\ e^{-r\theta}u_{i-1}\left(\theta_d^{i-1}, xe^{X_\theta} - D_i(xe^{X_\theta})\right)\mathbf{1}_{\{\tau=\theta\}} \end{array}\right] \tag{1.19}$$

où le supremum est pris sur tous les temps d'arrêt de la filtration générée par X qui sont $\mathbb{P}$-presque sûrement dans $[0, \theta]$.

On obtient alors $P(x) = u_I(t_d^I, x)$.

Frontière d'exercice

Nous allons nous intéresser aux fonctions intermédiaires. On fait encore une fois l'hypothèse maintenant qu'au moins une des conditions suivantes est vérifiée :

a) $\sigma > 0$,

b) $\int_{(0,1]} |x|\,\nu(dx) = +\infty$,

c) $\nu(\mathbb{R}_-^\star) > 0$.

Alors, nous pouvons énoncer le résultat suivant.

Proposition 1.2.7 *Pour $i \in \{0, \dots, I\}$, la fonction u_i est bornée et continue. De plus, à $\theta \geq 0$ fixé, la fonction $x \mapsto u_i(\theta, x)$ est décroissante et 1-Lispchitz. A l'exception du cas $i = 0$ et $\theta = 0$, u_i est strictement positive et il existe une fonction semi-continue supérieurement $c_i : \mathbb{R}_+ \to [0, K]$ telle que pour tout $\theta \geq 0$ $u_i(\theta, x) > (K - x)^+ \Leftrightarrow x > c_i(\theta)$.*

On obtient ainsi que le détenteur d'un Put Américain de maturité T et de strike K sur un tel sous jacent doit exercer au premier instant $t \in [0, T]$ où le prix de l'actif passe sous la frontière

$$t \mapsto c_I(t_d^I - t)\mathbf{1}_{\{t<t_d^I\}} + \sum_{i=0}^{I-1} c_i\left(\theta_d^i - (t - t_d^{i+1})\right)\mathbf{1}_{\{t\in[t_d^{i+1}, t_d^i)\}} + K\mathbf{1}_{\{t=T\}}. \tag{1.20}$$

Si on fait l'hypothèse légèrement plus restrictive maintenant qu'au moins une des conditions suivantes est vérifiée :

a) $\sigma > 0$,

b) $\int_{(0,1]} |x| \, \nu(dx) = +\infty$,

c) $\nu(\mathbb{R}_-^\star) > 0$ et le support de la mesure des sauts négatifs du processus de Lévy X contient 0.

Alors, on peut établir le résultat suivant.

Proposition 1.2.8 *Pour tout $i \in \{0, \ldots, I\}$, sous l'hypothèse que $\int_{\mathbb{R}}(e^y - 1)^+ \nu(dy) \leq r$, la fonction $c_i : \mathbb{R}_+ \mapsto [0, K]$ est continue.*

Cas Black-Scholes

Nous énonçons ici uniquement des résultats de [JJ12] établis dans le cadre du Chapitre 3 où le processus X est un Brownien drifté.

Proposition 1.2.9 (Smooth-fit) *Pour tout $i \in \{0, \ldots, I\}$, et pour tout $\theta > 0$, la fonction $x \mapsto u_i(\theta, x)$ est continûment dérivable.*

Nous donnons quelques éléments de preuve. On obtient ce résultat grâce à des contrôles assez fins sur les dérivées en temps de la fonction valeur. Ceux-ci se déduisent de la propriété de scaling du mouvement brownien. En combinant alors ces contrôles avec un argument classique de la théorie des solutions de viscosité, nous pouvons conclure.

Nous obtenons également des conditions de monotonie locale de la frontière au voisinage des instants de dividendes. Ces conditions peuvent s'exprimer uniquement à partir des dates et des fonctions de dividendes. Nous énonçons pour finir cette introduction un résultat sur le comportement de la frontière au voisinage des instants de dividendes.

Proposition 1.2.10 *Dans le modèle du Chapitre 3, si $i \geq 1$ et $0 < c_i(0) < c_{i-1}(\theta_d^{i-1})$, et si de plus $\liminf_{x \downarrow c_i(0)} \frac{D_i(x)}{x - c_i(0)} > 0$ alors on a :*

$$\lim_{\theta \to 0} \frac{c_i(\theta) - c_i(0)}{\sigma c_i(0) \sqrt{\theta \, |\ln \theta|}} = 1. \tag{1.21}$$

Organisation de la Partie I

Nous présenterons d'abord dans le Chapitre 3 les résultats dans le cas spécifique où la dynamique du sous-jacent entre les instants de dividendes est donnée par un modèle de Black-Scholes. Certains résultats ne sont vrais que dans ce cas particulier tandis que d'autres peuvent être généralisés au modèle exponentiel de Lévy. Nous énoncerons ces résultats plus généraux dans le Chapitre 4. Nous donnerons aussi des résultats liés à des choix particuliers de processus de Lévy.

2

Principe de programmation dynamique sous contraintes en probabilités

L'objectif de ce Chapitre est d'introduire les équations de programmation dynamique pour des problèmes de commande optimale en temps discret avec des contraintes spécifiques. Dans la Section 2.1, nous redonnerons l'énoncé du *principe de la programmation dynamique*, et montrerons qu'il permet d'établir des équations de programmation dynamique. Nous énoncerons des résultats à temps discret dans un cadre Markovien. Afin de faire le parallèle avec certains problèmes récents de contrôle optimal stochastique en temps continu, nous donnerons des résultats dans ce cadre. Dans la Section 2.2, nous donnerons des résultats sur les problèmes d'optimisation stochastique avec contraintes en espérance. Dans la sous-section 2.2.2, nous présenterons les résultats que nous avons obtenus dans la Partie II à propos de la résolution par programmation dynamique de problèmes de commande optimale discrets avec contraintes en probabilités.

2.1 Problème d'optimisation stochastique séquentiel

Comme énoncé en avant-propos, un problème d'optimisation stochastique fait intervenir un décideur qui souhaite minimiser un coût. Avant d'énoncer le principe d'optimalité de Bellman, nous allons préciser le type de problèmes d'optimisation stochastique que nous considérons. Puis, nous énoncerons des résultats généraux en temps discret. Enfin, nous particulariserons nos énoncés dans le but de pouvoir faire le lien avec la Section 2.2.

Pour une introduction plus détaillée sur la classification des problèmes d'optimisation stochastique, nous invitons le lecteur à consulter l'introduction de la thèse de Girardeau [Gir10].

2.1.1 Cadre de travail et principe d'optimalité de Bellman

Nous nous plaçons dans un cadre séquentiel. Ainsi, le temps est un ensemble ordonné discret noté $\mathbb{T}$ dont les éléments sont des dates notées génériquement t.

Dans toute la suite de cette introduction, une politique de décision est notée par $\mathbf{U}$. En particulier la décision prise à la date t est notée par $\mathbf{U}_t$. L'ensemble des décisions accessibles à la date t est noté $\mathbb{U}_t$. De manière plus générale, une variable aléatoire $\mathbf{Y}_t$ prend ses valeurs dans $\mathbb{Y}_t$.

Nous notons $(\mathcal{F}_t)_{t\in\mathbb{T}}$, la filtration accessible au décideur. Afin de simplifier les énoncés de cette introduction, nous ferons l'hypothèse que cette filtration est générée par une suite de variables aléatoires indépendantes $(\mathbf{W}_t)_{t\in\mathbb{T}}$. $\mathbf{W}_t$ prend ses valeurs dans $\mathbb{W}_t$, qui est muni d'une tribu naturelle. Nous notons alors $\mathbb{P}_t$ la loi de $\mathbf{W}_t$.

De manière systématique, nous noterons $\mathbb{Y}_{t_0:t_1}$ le produit cartésien des espaces $\mathbb{Y}_s$ pour s entre t_0 et t_1, i.e $\mathbb{Y}_{t_0:t_1} = \prod_{s=t_0}^{t_1} \mathbb{Y}_s$. Et de manière abusive, nous noterons $\mathbb{Y}$ au lieu de $\mathbb{Y}_{0:T}$.

Puisque le décideur ne peut pas prévoir l'avenir, ses décisions doivent être adaptées à la filtration $(\mathcal{F}_t)_{t\in\mathbb{T}}$. On notera $\mathcal{U}^{\mathrm{a}}$ l'ensemble des politiques de décision adaptées.

Sauf quand nous le préciserons explicitement, pour une fonction h de $\mathbb{Y}_1$ dans $\mathbb{Y}_2$ et pour une variable aléatoire $\mathbf{Y}_1$, la notation $h(\mathbf{Y}_1)$ est à comprendre comme la composition de h avec $\mathbf{Y}_1$ vu comme une fonction mesurable. Ainsi, si tout est compatible, $h(\mathbf{Y}_1)$ est une variable aléatoire à valeurs dans $\mathbb{Y}_2$.

Dans la suite, si nous ne précisons rien sur les espaces, c'est qu'il s'agit d'espaces topologiques séparés, leur tribu naturelle sera la tribu Borélienne.

Afin de caractériser des situations inacceptables pour le décideur, ou qui briseraient des contraintes d'état imposées par le problème, nous autorisons les fonctions réelles à prendre la valeur $+\infty$, comme il est classique en optimisation avec la règle d'arithmétique étendue donnée par $x \in \mathbb{R} \cup \{+\infty\}$, $x + (+\infty) = +\infty$. Ainsi le *coût* pour le décideur est une fonction φ de $\mathbb{U} \times \mathbb{W}$ dans $\mathbb{R} \cup \{+\infty\}$.

On note $\mathcal{U}^{\mathrm{a}}$ l'ensemble des contrôles qui sont adaptés à la filtration $(\mathcal{F}_t)_{t\in\mathbb{T}}$, et qui sont à valeurs dans $\mathbb{U}_t$ à la date t.

Le problème formalisé est donc pour le décideur de trouver un contrôle $\mathbf{U}^\sharp \in \mathcal{U}^{\mathrm{a}}$ tel que :

$$\mathbb{E}\left[\varphi\left(\mathbf{U}^\sharp, \mathbf{W}\right)\right] = \inf_{\mathbf{U}\in\mathcal{U}^{\mathrm{a}}} \mathbb{E}\left[\varphi\left(\mathbf{U}, \mathbf{W}\right)\right] \tag{2.1}$$

N'importe quelle politique de décision $\mathbf{U}^\sharp$ qui satisfait l'Equation (2.1) est une politique *optimale*.

En se plaçant sur un espace fonctionnel adapté, on pourrait probablement résoudre ce problème en utilisant une approche variationnelle liée au principe du minimum de Pontryagin (voir les thèses de

Girardeau [Gir10] et Barty [Bar04]). Cependant, nous allons préférer l'approche de Bellman [Bel54]. L'idée du *principe d'optimalité de Bellman* est la suivante :

Theorem 2.1.1 ([Bel54], Principe d'optimalité de Bellman) *An* optimal policy *has the property that no matter what the previous decisions (i.e controls) have been, the remaining decisions must constitute an optimal policy with regard to the state resulting from those previous decisions.*

De manière plus intuitive, en reprenant l'exemple du début de la Section 1.3 de Bertsekas [Ber01], le principe d'optimalité de Bellman dit que si le trajet optimal pour aller de Paris à Marseille passe par Lyon, alors arrivé à Lyon, le trajet optimal pour aller de Lyon à Marseille est de continuer sur le trajet initialement fixé.

L'idée du principe d'optimalité de Bellman se traduit en un principe dit *principe de programmation dynamique*. Ce principe définit un algorithme qui permet peut-être de résoudre le problème d'optimisation (2.1).

Définition 2.1. *L'*algorithme de la programmation dynamique *au sens d'Evstigneev [Evs76] se définit par récurrence rétrograde. Pour $t = T$, on pose :*

$$\varphi_T : u_{0:T-1} \times w_{0:T} \in \mathbb{U} \times \mathbb{W} \mapsto \varphi(u_{0:T}, w_{0:T}) \in \overline{\mathbb{R}}_+ \ . \tag{2.2a}$$

Puis pour $0 \leq t \leq T-1$, on pose :

$$\tilde{\varphi}_t : (u_{0:t}, w_{0:t}) \in \mathbb{U}_{0:t} \times \mathbb{W}_{0:t} \mapsto \mathbb{E}\left[\varphi_{t+1}\left(u_{0:t}, (w_{0:t}, \mathbf{W}_{t+1})\right)\right] \ , \tag{2.2b}$$

$$\varphi_t : (u_{0:t-1}, w_{0:t}) \in \mathbb{U}_{0:t-1} \times \mathbb{W}_{0:t} \mapsto \inf_{u_t \in \mathbb{U}_t} \tilde{\varphi}_t\left(u_{0:t+1}, w_{0:t+1}\right) \ . \tag{2.2c}$$

En toute généralité, rien ne garantit que l'algorithme théorique défini par (2.2) donne la stratégie optimale ainsi que la bonne valeur du problème d'optimisation (2.1). Cependant, nous pouvons donner des conditions suffisantes pour que cela soit vrai. Ces conditions précisées dans la définition 2.2 sont importantes dans le Chapitre 5.

Définition 2.2. *Si $\mathbb{U}$ est un espace métrique. Une fonction $\varphi : \mathbb{U} \times \mathbb{W} \to \mathbb{R} \cup \{+\infty\}$ est dite* intégrande normale en u *si elle est mesurable de* $\mathrm{Bor}(\mathbb{U}) \otimes \mathrm{Bor}(\mathbb{W})$ *dans* $\mathrm{Bor}(\mathbb{R} \cup \{+\infty\})$*, si pour tout $c \in \mathbb{R}$, et $w \in \mathbb{W}$, l'ensemble $\{u \in \mathbb{U} \mid \varphi(u, w) \leq c\}$ est fermé. Si pour tout $c \in \mathbb{R}$ et tout $w \in \mathbb{W}$, l'ensemble $\{u \in \mathbb{U} \mid \varphi(u, w) \leq c\}$ est compact alors φ est dite* inf-compacte en u.

Nous sommes en mesure d'énoncer un résultat d'existence de la stratégie optimale.

Proposition 2.1.2 ([Evs76]) *Si pour tout $t \in \mathbb{T}$, l'espace $\mathbb{U}_t$ est polonais, et si la fonction de coût est inf-compacte en u et minorée par une fonction $q : \mathbb{W} \to \mathbb{R} \cup \{+\infty\}$ qui soit $\mathbb{P}$-intégrable, alors il existe une stratégie optimale $\mathbf{U}^\sharp$ au problème d'optimisation* (2.1) *construite à l'aide de l'algorithme défini par* (2.2)*. La valeur de ce problème est donnée par $\mathbb{E}\left[\varphi_0(\mathbf{W}_0)\right]$.*

2.1.2 Cadre Markovien et fonction de coût additive en temps

Afin de pouvoir faire le lien avec les résultats sur les problèmes d'optimisation stochastique à temps continu, nous introduisons une généralisation du modèle Borélien à horizon fini de Bertsekas et Shreve [BS78].

Le modèle discret

A chaque date $t \in \mathbb{T}$, un *signal* $\mathbf{X}_t$ est observé. L'espace des valeurs possibles de $\mathbf{X}_t$ est noté $\mathbb{X}_t$. Ce signal a une dynamique connue, c'est-à-dire que l'on connait pour chaque date t, le noyau de transition de t à $t+1$ de la chaîne $(\mathbf{X}_t)_{t\in\mathbb{T}}$. On fait l'hypothèse qu'il existe une fonction mesurable g_{t+1} de $\mathbb{X}_t \times \mathbb{U}_t \times \mathbb{W}_{t+1}$ dans $\mathbb{X}_{t+1}$ telle que le noyau de transition de t à $t+1$ est donné par :

$$\mathbf{X}_{t+1} \overset{\mathcal{L}}{=} g_{t+1}(\mathbf{X}_t, \mathbf{U}_t, \mathbf{W}_{t+1}) \tag{2.3}$$

où $\overset{\mathcal{L}}{=}$ représente l'égalité en loi. Le signal est donc *contrôlé* par la décision $\mathbf{U}$ (dénommé aussi *contrôle*).

On fait l'hypothèse que la fonction de coût s'écrit comme une fonction déterministe des signaux, c'est-à-dire que le coût est une fonction φ de $\prod_{t=0}^{T} \mathbb{X}_t$ dans $\mathbb{R} \cup \{+\infty\}$.

Pour un contrôle adapté $\mathbf{U} \in \mathcal{U}^{\mathrm{a}}$, on note $\mathbf{X}^{\mathbf{U}}$ le signal contrôlé par $\mathbf{U}$, au sens où la dynamique de $\mathbf{X}$ est donnée par l'Equation (2.3) dans laquelle le contrôle $\mathbf{U}$ a été plongé.

Le problème formalisé est donc pour le décideur de trouver $\mathbf{U}^\sharp \in \mathcal{U}^{\mathrm{a}}$ tq :

$$\mathbb{E}\left[\varphi\left(\mathbf{X}^{\mathbf{U}^\sharp}\right)\right] = \inf_{\mathbf{U}\in\mathcal{U}^{\mathrm{a}}} \mathbb{E}\left[\varphi\left(\mathbf{X}^{\mathbf{U}}\right)\right] \tag{2.4}$$

Un problème du type de 2.4 est appelé processus de décision de Markov. Pour un traitement détaillé et de nombreux exemples, nous invitons le lecteur à consulter les ouvrages de Bauerle et Rieder [BR10], de Bertsekas [Ber01], ou de Puterman [Put94].

N'importe quelle politique de décision $\mathbf{U}^\sharp$ qui satisfait l'Equation (2.4) est une politique *optimale*. Il est à noter que la valeur $\mathbb{E}\left[\varphi\left(\mathbf{X}^{\mathbf{U}^\sharp}\right)\right]$ dépend de la loi de $\mathbf{X}_0$.

Pour toutes les dates t, les espaces $\mathbb{X}_t$ et $\mathbb{U}_t$ sont supposées être Polonais. De plus, nous considérons que la fonction coût du problème est donnée sous la forme d'une somme de fonctions simples. Pour être plus précis, nous faisons l'hypothèse qu'il existe une famille de fonctions $(L_t : \mathbb{X}_t \to \mathbb{R}_+ \cup \{+\infty\})_{t=0,\cdots,T-1}$, ainsi qu'une fonction terminale $K : \mathbb{X}_T \to \mathbb{R}_+ \cup \{+\infty\}$ telles que $\varphi(x_{0:T}) = \sum_{t=0}^{T-1} L_t(x_t) + K(x_T)$ où $x_{0:T}$ est une notation commode pour parler du $(T+1)$-uplet $(x_0, \cdots, x_T)$.

Sous cette hypothèse, le principe de la programmation dynamique énoncé par Bellman conduit à l'équation de Bellman.

Plus rigoureusement, pour $t = T$, on définit :

$$V_t : x_T \in \mathbb{X}_T \mapsto K(X_T) \tag{2.5a}$$

puis pour $t \in \{0, \cdots, T-1\}$, on définit :

$$V_t : x_t \in \mathbb{X}_t \mapsto \inf_{u_t \in \mathbb{U}_t} L_t(x_t) + \mathbb{E}\left[V_{t+1}\left(\mathbf{X}_{t+1}^{u_t}\right) \middle| \mathcal{F}_t, \mathbf{X}_t = x_t \right]. \tag{2.5b}$$

mais à cause du caractère Markovien, et de la dynamique particulière du signal contrôlé, on obtient :

$$V_t(x_t) = \inf_{u_t \in \mathbb{U}_t} L_t(x_t) + \mathbb{E}\left[V_{t+1}\left(g_{t+1}(x_t, u_t, \mathbf{W}_{t+1})\right)\right]. \tag{2.5c}$$

On suppose que les fonctions $(g_t)_{t=1,\ldots,T}$ sont globalement mesurables et continues partiellement en (x_t, u_t), et pour tout $t \in \{0, \cdots, T-1\}$, on introduit pour tout $u_t \in \mathbb{U}_t$, l'opérateur $\mathrm{P}_{t,t+1}^{u_t}$ qui à une fonction continue $f : \mathbb{X}_{t+1} \to \mathbb{R}_+ \cup \{+\infty\}$ associe une fonction continue $\mathrm{P}_{t,t+1}^{u_t} f : \mathbb{X}_t \to \mathbb{R}_+ \cup \{+\infty\}$ définie par :

$$\mathrm{P}_{t,t+1}^{u_t} f : x_t \in \mathbb{X}_t \mapsto \mathbb{E}\left[f\left(g_{t+1}(x_t, u_t, \mathbf{W}_{t+1})\right)\right]. \tag{2.6}$$

Il est à noter que la continuité globale $\mathrm{P}_{t,t+1}^{u_t} f$ vu comme fonction de (x_t, u_t) est une conséquence de la caractérisation séquentielle de la continuité dans les espaces métriques et du théorème de convergence dominée.

Nous obtenons ainsi que l'Equation (2.5c) se réécrit :

$$V_t(x_t) = L_t(x_t) + \inf_{u_t \in \mathbb{U}_t} \mathrm{P}_{t,t+1}^{u_t} V_{t+1}(x_t) \tag{2.7}$$

L'Equation (2.7) s'appelle l'équation de Bellman, on peut noter qu'elle est similaire à l'Equation (1.5), qui s'appelle Equation de Wald-Bellman et qui est un cas particulier où le contrôle est soit d'arrêter, soit de continuer.

Cependant, nous n'avons toujours pas établi si la fonction $\mathbb{E}\left[V_0(\mathbf{X_0})\right]$ est bien égale à la valeur du Problème (2.4) initialement posé, et nous n'avons pas non-plus établi si il existait des stratégies optimales.

Existence d'une politique optimale

Nous rappelons ici un résultat d'existence dans le cadre où on suppose que pour tout $t \in \{0, \cdots, T-1\}$, l'espace $\mathbb{U}_t$ est un espace métrique compact, L_t est une fonction semi-continue inférieurement et positive et g_t est une fonction continue.

Theorem 2.1.3 (Prop.8.6 [BS78]) *Sous les hypothèses précédentes, le problème* (2.4) *a une solution* $\mathbf{U}^\sharp$ *qui se construit à l'aide de l'équation de Bellman* (2.7)*. De plus, le contrôle optimal à la date t ne dépend que de l'état $\mathbf{X}_t$ à la date t.*

Nous tenons à souligner le fait que dans le cas précédent, le contrôle optimal à la date t ne dépend que de la valeur de l'état, et que ce contrôle est optimal dans toutes les stratégies qui seraient adaptées.

Bien que la démonstration du résultat précédent ne soit pas énoncée ici, nous faisons part que celle-ci est caractéristique de la manière de trouver le contrôle optimal et d'établir que l'équation de Bellman résoud le problème de minimisation initial. Il s'agit de montrer que la propriété de régularité nécessaire pour avoir l'existence d'un minimiseur se propage de la date $t+1$ à la date t.

L'énoncé en temps continu

Bien que ce ne soit pas le cadre direct de la thèse, nous rappelons ici l'énoncé du problème en temps continu afin de pouvoir mettre en valeur dans §2.2 la correspondance entre le temps continu et le temps discret dans la prise en compte des contraintes en espérance dans un problème en temps continu.

En temps continu, l'équation de Bellman devient l'équation d'Hamilton-Jacobi-Bellman. En effet, dans le cadre du temps continu, l'opérateur $\mathrm{P}^u_{t,t+1}$ (voir (2.6)) doit se comprendre $I + dt \times \mathrm{L}^u_t$ où L^u_t est le générateur infinitésimal du processus de Markov qui pendant un bref instant entre t et $t+dt$, aurait été contrôlé suivant le contrôle u. L'équation d'Hamilton-Jacobi-Bellman est donc de manière très informelle :

$$V_t(x_t) = V_{t+dt}(x_t) + dt \times \left(\inf_{u_t \in \mathbb{U}_t} L_t(x_t) + \mathrm{L}^{u_t}_t V_{t+dt}(x_t) \right) . \tag{2.8}$$

Nous énonçons ici un résultat rigoureux dans le cas où le signal est solution d'une équation différentielle stochastique contrôlée définie sur un espace de Wiener n-dimensionnel.

On note L^u l'opérateur infinitésimal d'un processus de diffusion $\mathbf{X}^{\mathbf{U}}$ contrôlé à valeurs dans $\mathbb{X} = \mathbb{R}^n$ pour un entier n. On suppose que, pour tous $u, u' \in \mathbb{U}$, $\mathrm{dom}(\mathrm{L}^u) = \mathrm{dom}(\mathrm{L}^{u'})$. On note $\mathrm{dom}(\mathrm{L}) = \cap_{u \in \mathbb{U}} \mathrm{dom}(\mathrm{L}^u)$. On fait l'hypothèse qu'il existe $b : \mathbb{X} \times \mathbb{U} \to \mathbb{R}^n$ et $\sigma : \mathbb{X} \times \mathbb{U} \to \mathbb{R}^{n \times n}$ telles que pour $f : \mathbb{X} \to \mathbb{R}$ appartenant à $\mathrm{dom}(\mathrm{L}^u)$ et $(x, u) \in \mathbb{X} \times \mathbb{U}$:

$$\mathrm{L}^u f(x) = b(x,u).\nabla f(x) + \frac{1}{2} \mathrm{Trace} \left(\sigma(t,x)\ {}^t\sigma(t,x) \nabla^2 f(x) \right) . \tag{2.9}$$

Pour tout contrôle $\mathbf{U}$ adapté, le processus $\mathbf{X}^{\mathbf{U}}$ est le processus adapté tel que pour tout $f \in \mathrm{dom}(\mathrm{L})$, $\lim_{h \to 0_+} \frac{1}{h} \mathbb{E} \left[f\left(\mathbf{X}^{\mathbf{U}}_{t+h}\right) - f\left(\mathbf{X}^{\mathbf{U}}_t\right) \middle| \mathcal{F}_t \right] = \mathrm{L}^{\mathbf{U}_t} f(\mathbf{X}_t)$.

Theorem 2.1.4 (Th.4.3.1, Rem.4.3.4 [Pha09]) *Soit* $l : (t, x, u) \in [0, T] \times \mathbb{X} \times \mathbb{U} \to \mathbb{R}_+$. *On s'intéresse à la fonction valeur* v *définie pour* $t \in [0, T]$ *et* $x \in \mathbb{X}$ *:*

$$v(t, x) = \inf_{\mathbf{U}\,adapté} \mathbb{E}\left[\int_t^T l\left(s, \mathbf{X}_s^{\mathbf{U}}, \mathbf{U}_s\right) ds \middle| \mathcal{F}_t, \mathbf{X}_t^{\mathbf{U}} = x \right] \tag{2.10}$$

Pour $(t, x, v) \in [0, T] \times \mathbb{X} \times \mathrm{dom}(\mathrm{L})$, *on pose* $H(t, x, v) = \inf_{u \in \mathbb{U}} \mathrm{L}^u v(x) + l(t, x, u)$. *Et on fait l'hypothèse que* $\mathrm{dom}(H) = [0, T] \times \mathbb{X} \times \mathrm{dom}(\mathrm{L})$. *Alors* v *est solution de viscosité de*

$$-\partial_t v(t, x) - H(t, x, v) = 0, \ (t, x) \in [0, T) \times \mathbb{X}. \tag{2.11}$$

L'équation (2.11) est une équation d'Hamilton-Jacobi-Bellman. Elle est l'énoncé rigoureux de l'équation formelle (2.8).

Pour un exposé plus précis des résultats sur le contrôle optimal en temps stochastique, le principe de programmation dynamique et les solutions de viscosité, nous renvoyons le lecteur au livre de Flemming et Soner [FS93] ainsi qu'au livre de Kushner et Dupuis [KD00] pour les implications numériques de ce genre de résultats.

2.2 Contrainte en espérances

Nous reprenons le modèle de la section précédente où le signal $\mathbf{X}^{\mathbf{U}}$ est contrôlé par $\mathbf{U}$, et où le coût à minimiser est donné par φ. Mais nous introduisons en plus une contrainte en espérance sous la forme générique que $\mathbb{E}\left[G(\mathbf{X}_T^{\mathbf{U}})\right] \leq p$ pour une certaine fonction $G : \mathbb{X}_T \to \mathbb{R}$ et un nombre réel p.

Le problème formalisé est donc pour le décideur de trouver $\mathbf{U}^\sharp \in \mathcal{U}^{\mathrm{a}}$ tel que $\mathbb{E}\left[G(\mathbf{X}_T^{\mathbf{U}^\sharp})\right] \leq p$:

$$\mathbb{E}\left[\varphi\left(\mathbf{X}^{\mathbf{U}^\sharp}\right)\right] = \inf_{\mathbf{U} \in \mathcal{U}^{\mathrm{a}} : \mathbb{E}\left[G(\mathbf{X}_T^{\mathbf{U}})\right] \leq p} \mathbb{E}\left[\varphi\left(\mathbf{X}^{\mathbf{U}}\right)\right] \tag{2.12}$$

2.2.1 Problèmes de contrôle avec cible stochastique

Initiés par Soner et Touzi [ST02b] pour des questions de surreplication de portefeuille d'actifs financiers, les problèmes de *cible stochastique* reposent très fortement sur l'idée de programmation dynamique [ST02a] et sur l'idée de rajouter un contrôle martingale pour prendre en compte la contrainte en espérance [BET10].

Plaçons nous dans un cadre de processus de diffusion contrôlé analogue à la §2.1.2. On introduit le problème (2.10) partant de t avec la contrainte en espérance $\mathbb{E}\left[G(\mathbf{X}_T^{\mathbf{U}}) | \mathcal{F}_t\right] \leq p$. I.e, nous définissons pour $(x, p) \in \mathbb{X} \times \mathbb{R}$:

$$\bar{v}(t,x,p) = \inf_{\substack{\mathbf{U}\in\mathcal{U}^{\mathrm{a}} \\ \mathbb{E}\left[G(\mathbf{X}_T^{\mathbf{U}})\middle|\mathbf{X}_t^{\mathbf{U}}=x\right]\geq p}} \mathbb{E}\left[\int_t^T l\left(s,\mathbf{X}_s^{\mathbf{U}},\mathbf{U}_s\right)ds\middle|\mathcal{F}_t,\mathbf{X}_t^{\mathbf{U}}=x\right] \tag{2.13}$$

alors nous savons que :

Proposition 2.2.1 (Prop.5.1 [BEI10]) *Le problème* (2.13) *est solution du problème de contrôle avec contrainte de cible stochastique suivant.*

On définit pour $(t,x,p)\in[0,T]\times\mathbb{X}\times\mathbb{R}$:

$$\bar{\mathcal{U}}_{t,x,p} := \left\{\begin{pmatrix}\mathbf{U}\\ \boldsymbol{\alpha}\end{pmatrix}\in\begin{pmatrix}\mathcal{U}^{\mathrm{a}}\\ \mathcal{A}^{\mathrm{a}}_{2-integrable}\end{pmatrix},\ G\left(\mathbf{X}_T^{\mathbf{U}}|_{\mathbf{X}_t^{\mathbf{U}}=x}\right)-p-\int_t^T\boldsymbol{\alpha}_s.dW_s\leq 0\right\} \tag{2.14a}$$

où W est le brownien sous-jacent à partir duquel sont construits les solutions fortes de la diffusion contrôlée, et $\mathcal{A}^{\mathrm{a}}_{2-integrable}$ est l'ensemble des processus adaptés à valeurs dans $\mathbb{R}^n$ qui sont de carré intégrable sur $[0,T]\times\Omega$ pour la mesure $dt\otimes d\mathbb{P}$. On note génériquement $\bar{\mathbf{U}}$ la paire $\mathbf{U},\boldsymbol{\alpha}$ et on a :

$$\bar{v}(t,x,p) = \inf_{\bar{\mathbf{U}}\in\bar{\mathcal{U}}_{t,x,p}}\mathbb{E}\left[\int_t^T f(\mathbf{X}_s^{\mathbf{U}})ds|\mathbf{X}_t^{\mathbf{U}}=x\right]\ . \tag{2.14b}$$

A la manière de la proposition précédente, nous énonçons maintenant en quoi le problème se ramène à un problème de décision Markovien.

Proposition 2.2.2 (Th.5.5 [Gir10]) *Il y a équivalence entre l'existence d'une solution pour le problème* (2.12) *et l'existence d'une solution pour le problème suivant :*

$$\inf_{\substack{\mathbf{U}\in\mathcal{U}^{\mathrm{a}},\\ \mathbf{V}\in\mathcal{V}^{\mathrm{a}},\\ G(\mathbf{X}_T)-\mathbf{Z}_T\leq p}} \mathbb{E}\left[\varphi\left(\mathbf{X}^{\mathbf{U}}\right)\right] \tag{2.15a}$$

où $\mathcal{V}^{\mathrm{a}}$ est l'ensemble des processus réels adaptés tels que $\mathbb{E}\left[\mathbf{V}_{t+1}|\mathcal{F}_t\right]\leq 0$, et où $\mathbf{Z}^{\mathbf{V}}$ (noté $\mathbf{Z}$) est le processus contrôlé réel de valeur initiale nulle tel que :

$$\mathbf{Z}_{t+1} = \mathbf{Z}_t + \mathbf{V}_{t+1}. \tag{2.15b}$$

Dans le problème de Bouchard, Elie et Touzi [BET10], afin de pouvoir écrire un principe de programmation dynamique, les auteurs ajoutent un contrôle martingale qui permet de transformer les contraintes en espérances en des contraintes presque-sûres. La Proposition précédente en est l'analogue à temps discret. En suivant [CCC$^+$11], on peut montrer que, pour le problème (2.15), il est naturel de dériver (au moins formellement) une équation de Bellman étendue, en posant

$$V_T : (x_{0:T}, p) \in \mathbb{X}_{0:T} \times \mathbb{R} \mapsto \begin{cases} \varphi(x_{0:T}) & \text{si } G(x_T) \leq p \\ +\infty & \text{sinon} \end{cases} \tag{2.16a}$$

et les problèmes pour $t = 0, \cdots, T-1$:

$$V_t : (x_{0:t}, p) \in \mathbb{X}_{0:t} \times \mathbb{R} \mapsto \inf_{\substack{u_t \in \mathbb{U}_t, \\ \mathbf{V}_{t+1} : \mathbb{E}[\mathbf{V}_{t+1} | \mathcal{F}_t] \leq 0}} \mathbb{E}\left[V_{t+1}\left(x_{0:t}, g_{t+1}(x_t, u_t, \mathbf{W}_{t+1}), p + \mathbf{V}_{t+1}\right)\right]. \tag{2.16b}$$

Cependant, il n'existait pas à notre connaissance de résultat permettant d'affirmer que $\mathbb{E}\left[V_0(\mathbf{X}_0)\right]$ est la solution du problème d'optimisation (2.12). La réponse à cette question est l'une des motivations de la Partie II.

2.2.2 Résumé des résultats de la Partie II

Dans la Partie II, nous généralisons l'idée de la programmation dynamique au sens d'Evstigneev [Evs76] ainsi que le résultat de la Proposition 2.1.2 afin d'obtenir un principe de programmation dynamique avec des contraintes en espérance conditionnelle. Au coeur de cette généralisation, se trouve le concept de relaxation de Young [You37]. Il s'agit de plonger l'espace $L^0(\mathbb{W}, \mathbb{U})$ des fonctions "mesurables" de $\mathbb{W}$ dans $\mathbb{U}$ dans l'espace $\mathcal{R}(\mathbb{W}, \mathbb{U})$ des fonctions de $\mathbb{W}$ dans $\mathcal{M}_1(\mathbb{U})$ (l'espace des probabilités sur $\mathbb{U}$). En supposant que $\mathbb{W}$ et $\mathbb{U}$ sont des espaces topologiques appropriés, on équipe ensuite cet espace d'une topologie affaiblie adaptée à notre cadre. Cette approche issue des problèmes de calcul variationnel et mis en valeur par Berliocchi et Lasry [BL73] a été appliquée pour des problèmes d'optimisation stochastique par Balder [Bal84, Bal00], mais aussi par Pedregal [Ped97, Ped99]. Pour un traitement plus récent et exhaustif du bon cadre topologique des mesures de Young, nous renvoyons le lecteur à l'excellent ouvrage de Castaing, Raynaud de Fitte et Valadier [CRdFV04].

On se replace dans le cadre général du problème d'optimisation (2.1), à ceci près qu'il existe désormais un contrôle terminal $\mathbf{U}_T$ à la date T. On fait les mêmes hypothèses que dans la proposition 2.1.2 sur la fonction φ et les espaces $\mathbb{U}_t$ pour $t \in \mathbb{T}$. On fait l'hypothèse supplémentaire que les espaces $\mathbb{W}_t$ sont des espaces de Lusin pour tout $t \in \mathbb{T}$.

On se donne une famille de fonctions $(\gamma_t)_{t \in \mathbb{T}}$ telle que pour tout $t \in \mathbb{T}$, $\gamma_t : \mathbb{U}_{0:t} \times \mathbb{W}_{0:t} \to \mathbb{R} \cup \{+\infty\}$ est une intégrande normale en $u_{0:t}$ (voir Définition 2.2). Pour simplifier les énoncés de cette introduction, on supposera pour tout $t \in \mathbb{T}$ que γ_t est minorée inférieurement.

On définit alors $\mathcal{U}^\gamma$ l'ensemble des contrôles $\mathbf{U} = (\mathbf{U}_t)_{t \in \mathbb{T}}$ tel que pour tout $t \in \mathbb{T}, t \neq T$, $\mathbb{E}\left[\gamma_t(\mathbf{U}_{0:t+1}, \mathbf{W}_{0:t+1}) | \mathcal{F}_t\right] \leq 1$ $\mathbb{P}$-presque sûrement.

Le nouveau problème que l'on considère est de trouver $\mathbf{U}^\sharp \in \mathcal{U}^{\mathrm{a}} \cap \mathcal{U}^\gamma$ tel que :

$$\mathbb{E}\left[\varphi\left(\mathbf{U}^\sharp, \mathbf{W}\right)\right] = \inf_{\mathbf{U} \in \mathcal{U}^{\mathrm{a}} \cap \mathcal{U}^\gamma} \mathbb{E}\left[\varphi\left(\mathbf{U}, \mathbf{W}\right)\right] \tag{2.17}$$

On définit alors un nouvel algorithme de programmation dynamique dans l'esprit de celui défini par (2.2). Pour $t = T + 1$, on pose $\bar{\varphi}_{T+1} \equiv \varphi$, puis pour $0 \leq t \leq T$, on définit l'espace des mesures de Young $\mathcal{R}(\mathbb{W}_t, \mathbb{U}_t)$ dont les éléments sont des fonctions de $\mathbb{W}_t$ dans l'espace des probabilités sur $(\mathbb{U}_t, \mathrm{Bor}(\mathbb{U}_t))$ puis on définit les fonctions suivantes $I^{\mathbb{P}_t}_{\gamma_{t-1}} : \mathbb{U}_{0:t-1} \times \mathbb{W}_{0:t-1} \to \mathbb{R} \cup \{+\infty\}$ tel que :

$$I^{\mathbb{P}_t}_{\gamma_{t-1}} : (u_{0:t-1}, w_{0:t-1}, \mu_t) = \int_{\mathbb{W}_t} \int_{\mathbb{U}_t} \gamma_{t-1}(u_{0:t}, w_{0:t}) \mu_t(w_t)(du_t) \mathbb{P}_t(dw_t) \tag{2.18a}$$

ainsi que les ensembles suivants :

$$\mathcal{R}^{\gamma_{t-1}}(u_{0:t-1}, w_{0:t-1}) = \left\{ \mu_t \in \mathcal{R}(\mathbb{W}_t, \mathbb{U}_t) \;\middle|\; I^{\mathbb{P}_t}_{\gamma_{t-1}}(u_{0:t-1}, w_{0:t-1}, \mu_t) \leq 1 \right\} \tag{2.18b}$$

et avec la convention que lorsque $t = 0$, $\mathcal{R}^{\gamma_{t-1}} = \mathcal{R}(\mathbb{W}_0, \mathbb{U}_0)$. Enfin, on définit la suite de fonctions $\bar{\varphi}_t : \mathbb{U}_{0:t-1} \times \mathbb{W}_{0:t-1} \to \mathbb{R} \cup \{+\infty\}$ tel que :

$$\bar{\varphi}_t : (u_{0:t-1}, w_{0:t-1}) = \inf_{\mu_t \in \mathcal{R}^{\gamma_{t-1}}(u_{0:t-1}, w_{0:t-1})} \int_{\mathbb{W}_t} \int_{\mathbb{U}_t} \bar{\varphi}_{t+1}(u_{0:t}, w_{0:t}) \mu_t(w_t)(du_t) \mathbb{P}_t(dw_t) \; . \tag{2.18c}$$

On obtient alors le résultat suivant :

Proposition 2.2.3 *Sous les hypothèses précédentes, et l'hypothèse supplémentaire que pour tout $t \in \mathbb{T}$, $(\mathbb{W}_t, \mathrm{Bor}(\mathbb{W}_t), \mathbb{P}_t)$ est un espace de probabilité sans atomes, alors le problème de minimisation* (2.17) *admet une solution, qui peut être déterminée par l'algorithme de programmation dynamique* (2.18)*. Sa valeur est donnée par $\bar{\varphi}_0$.*

Comme cas particulier, nous prouvons, à propos du problème de minimisation (2.15), le résultat suivant.

Proposition 2.2.4 *Sous les hypothèses précédentes, et l'hypothèse supplémentaire que pour tout $t \in \mathbb{T}$, $(\mathbb{W}_t, \mathrm{Bor}(\mathbb{W}_t), \mathbb{P}_t)$ est un espace de probabilité sans atomes, et celles de la sous-sous-section 2.1.2 portant sur les espaces $\mathbb{U}_t$ et la fonction de coût, et si la fonction G est semi-continue inférieurement et bornée inférieurement, alors les équations de Bellman étendues* (2.16) *résolvent le problème* (2.15)*.*

Part I

American contingent claims with discrete dividends

3

Regularity of the American Put option in the Black-Scholes model with general discrete dividends

Summary. We analyze the regularity of the value function and of the optimal exercise boundary of the American Put option when the underlying asset pays a discrete dividend at known times during the lifetime of the option. The ex-dividend asset price process is assumed to follow the Black-Scholes dynamics and the dividend amount is a deterministic function of the ex-dividend asset price just before the dividend date. This function is assumed to be non-negative, non-decreasing and with growth rate not greater than 1. We prove that the exercise boundary is continuous and that the smooth contact property holds for the value function at any time but the dividend dates. We thus extend and generalize the results obtained in [JV11] when the dividend function is also positive and concave. Lastly, we give conditions on the dividend function ensuring that the exercise boundary is locally monotonic in a neighborhood of the corresponding dividend date.

Introduction

We consider the American Put option with maturity T and strike K written on an underlying stock S. Like in [JV11], we assume that the stochastic dynamics of the ex-dividend price process of this stock can be modelled by the Black Scholes model and that this stock is paying discrete dividends at deterministic times $0 \leq t_d^I < t_d^{I-1} < \cdots < t_d^i < \cdots < t_d^1 < T$. At each dividend time t_d^i, the value of the stock becomes $S_{t_d^i} = S_{t_d^i-} - D_i\left(S_{t_d^i-}\right)$ where $D_i(S_{t_d^i-})$ is the value of the dividend payment (see Figure 3.1). We suppose that each dividend function $D_i : \mathbb{R}_+ \to \mathbb{R}_+$ is non-decreasing, non-negative and such that $x \mapsto x - D_i(x)$ is also non-decreasing and non-negative.
We are interested in the value of the American Put option with strike K and maturity T. Since we are in a Markovian framework, the price can be characterized in terms of a value function depending of the time t and the stock price at time t. For the sake of consistency, we will denote this value function by u_0 for the case without dividends.

By change of numeraire, the pricing problem of the American Put option in the Black-Scholes model with continuously paid proportional dividends is equivalent to the pricing problem of the Amer-

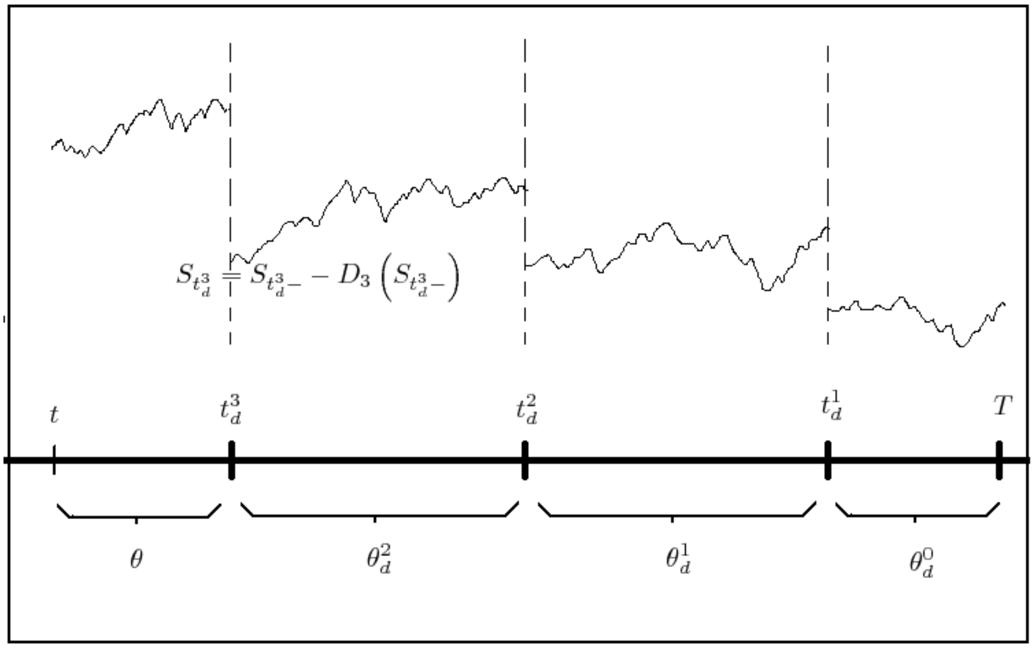


Fig. 3.1. A trajectory of the stock price process

ican Call option obtained by exchange of the spot value of the underlying and the strike and exchange of the dividend and interest rates. The latter problem was studied in [McK65] by McKean who first linked this optimal stopping-time problem to a free boundary problem involving both the value function and the exercise boundary. Van Moerbeke [VM76] derived an integral equation which involves both the exercise boundary and its derivative. Kim [Kim90] later obtained an integral equation which only involves the exercise boundary itself. Independently, Jacka [Jac93] and Carr, Jarrow and Myneni [CJM92] derived the analogue equation for the exercise boundary c_0 of the American Put option in the Black-Scholes model without dividends. According to Jacka [Jac93], the boundary c_0 is continuous, the first time the price process crosses c_0 is an optimal stopping time and the smooth fit property holds for the value function u_0. The uniqueness for the integral equation was left as an open problem in those papers. Uniqueness was proven by Peskir [Pes05]. We refer to [PS06, Section.25.] for a more recent exposition of these results. Convexity of c_0 was proved in [CCJZ08] and in [Eks04]. Infinite regularity of c_0 at all points prior to the maturity was formally proved by Chen and Chadam [CC06]. Then Bayraktar and Xing [BX09] proved that this remains true if the underlying asset pays continuous dividends at a fixed rate. In practice, continuous dividends are not a satisfying model since dividends are paid once a year or quarterly. That is why we are interested in dividends that are paid at a number of discrete points in time.

When we assume discrete dividend payments, in general, the value function of the Put option will no longer be convex in the stock price variable, even if convexity is preserved for linear dividend functions. Moreover, the optimal exercise boundary will become discontinuous at the dividend dates and before the dividend dates it may not be monotone. Integral formulas for the exercise boundary which are similar to the ones in [CJM92] have been derived under the assumption that the boundary is Lipschitz continuous (see Göttsche and Vellekoop [GV11]) or locally monotonic (Vellekoop & Nieuwenhuis [VN11]). In this paper we continue the study, undertaken in [JV11], of conditions under which such regularity properties of the optimal exercise boundary under discrete dividend payments can be proven.

We prove that the exercise boundary is continuous at any time which is not a dividend date and that the smooth contact property holds for the value function of the option. We considerably extend the results obtained in [JV11], where the continuity of the exercise boundary and the smooth contact property were only obtained in a left-hand neighborhood of the first dividend date when the corresponding dividend function was assumed to be globally concave and linear with a positive slope in a neighbohood of the origin. Under the much more restrictive assumption of global linearity of all the dividend functions, the smooth contact property and the right-continuity (resp. continuity) of the exercise boundary was proved to hold globally (resp. in a left-hand neighborhood of each dividend date). We also extend the result obtained in [JV11] on the decrease of the exercise boundary in a left-hand neighborhood of the first (resp of each) dividend date when the corresponding dividend function was assumed to be positive and concave (resp. when all dividend functions were supposed to be linear) : we give more general sufficient conditions on each dividend function for the exercise boundary to be either non-decreasing or non-increasing in a left-hand neighborhood of the corresponding dividend date.

In the first section, we introduce our notations and assumptions. In the second section, we recall the existence results for the value function and the exercise boundary stated in [JV11]. The third section is devoted to the smooth-fit property and relies on a viscosity solution approach combined with an estimation of the derivative of the value function with respect to the time variable. In the fourth section, we prove the continuity result for the exercise boundary, which is known to be upper-semicontinuous by continuity of the value function. The right-continuity is obtained by comparison with the optimal boundary of the Put option in the Black-Scholes model without dividend. The left-continuity follows from the characterization of the continuation region as the set of points where the spatial derivative of the value function is greater than -1. In the fifth section, we are interested in the local behaviour of the exercise boundary in a neighborhood of the dividend date. To be able to analyse this behaviour, we have to assume that the stock level at which the dividend function becomes positive lies in the post-dividend exercise region. When the dividend function has a positive slope at this point, we obtain a first order expansion for the exercise boundary at the dividend date. We also provide sufficient conditions for the exercise boundary to be locally monotonic.

3.1 Notations and assumptions

3.1.1 Notations

- $\left(\Omega, \mathcal{F}, (\mathcal{F}_s)_{s\geq 0}, \mathbb{P}\right)$ is a probability space with a right continuous filtration, $(B_s)_{s\geq 0}$ a $(\mathcal{F}_s)$-Brownian motion under $\mathbb{P}$, and $\mathbb{Q}$ is the probability measure defined by
$$\left.\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}\right|_{\mathcal{F}_t} = e^{-\frac{\sigma^2}{2}t+\sigma B_t}.$$
- $\bar{S}^x_t$ is a geometric Brownian motion satisfying : $d\bar{S}^x_t = r\bar{S}^x_t dt + \sigma \bar{S}^x_t dB_t$ and $\bar{S}^x_0 = x$. Its density at time t is denoted $\mathrm{p}(t,x,y) = \dfrac{\mathbf{1}_{\{y>0\}}}{\sigma y\sqrt{2\pi t}} \exp\left(-\frac{1}{2\sigma^2 t}\left(\ln\left(\frac{y}{x}\right) - (r-\frac{\sigma^2}{2})t\right)^2\right)$,
- $\mathcal{A}$ is the Black-Scholes operator defined for any $\mathcal{C}^2$ function f by $\mathcal{A}f(x) = -rf(x) + rxf'(x) + \frac{\sigma^2 x^2}{2}f''(x)$,
- the set of all the stopping times of $(\mathcal{F}_s)_{s\leq\theta}$ is abusively denoted by $\{\tau \in [0,\theta]\}$.

Recursive construction

Let $(\theta^i_d = t^i_d - t^{i+1}_d)_{0\leq i\leq I-1}$ with the convention $t^0_d = T$ denote the durations between the dividend dates. For non-negative values of θ and x, we define by induction

- $u_0(\theta,x) = \sup_{\tau\in[0,\theta]} \mathbb{E}\left[e^{-r\tau}\left(K-\bar{S}^x_\tau\right)^+\right]$ the price of the American Put option in the Black-Scholes model without dividends when the time to maturity is θ and the spot level x. The corresponding exercise boundary is $c_0(\theta)$ such that $\{x : u(0,x) > (K-x)^+\} = (c_0(\theta), +\infty)$. Let $v(\theta,x)$ be the value function of the American Put option with normalized strike 1 in the Black Scholes model without dividends and $\bar{c}(\theta)$ the associated exercise boundary. One has :
$$u_0(\theta,x) = \sup_{\tau\in[0,\theta]} \mathbb{E}\left[e^{-r\tau}\left(K-\bar{S}^x_\tau\right)^+\right] = K\sup_{\tau\in[0,\theta]} \mathbb{E}\left[e^{-r\tau}\left(1-\bar{S}^{x/K}_\tau\right)^+\right] = Kv\left(\theta, \frac{x}{K}\right)$$
and consequently $c_0(\theta) = \sup\{x | u_0(\theta,x) = (K-x)^+\} = K\bar{c}(\theta)$.
- $\forall i \in \{1,\ldots,I\}$,
$$u_i(\theta,x) = \sup_{\tau\in[0,\theta]} \mathbb{E}\left[e^{-r\tau}\left(K-\bar{S}^x_\tau\right)^+ \mathbf{1}_{\{\tau<\theta\}} + e^{-r\theta}u_{i-1}(\theta^{i-1}_d, \bar{S}^x_\theta - D_i(\bar{S}^x_\theta))\mathbf{1}_{\{\tau=\theta\}}\right].$$
Note that $u_i(0,x) = u_{i-1}(\theta^{i-1}_d, x - D_i(x))$.
- Any stopping time τ such that $u_i(\theta,x) = \mathbb{E}\left[e^{-r\tau}\left(K-\bar{S}^x_\tau\right)^+ \mathbf{1}_{\{\tau<\theta\}} + e^{-r\theta}u_i(0,\bar{S}^x_\theta)\mathbf{1}_{\{\tau=\theta\}}\right]$ will be abusively called **an optimal stopping time for** $u_i(\theta,x)$.

3.1.2 Assumptions

In all what follows, we assume that

$$(A)\ \forall i \in \{1,\dots,I\}, \begin{cases} \text{(a) } D_i \text{ is non-decreasing and non-negative,} \\ \text{(b) } \rho_i : x \mapsto x - D_i(x) \text{ is non-decreasing and non-negative.} \end{cases}$$

3.2 Previous results

Under (A), we can reformulate Proposition 1.5 [JV11] with our notations,

Proposition 3.2.1 *Suppose that $t < t_d^i < t_d^{i-1} < \cdots < t_d^1 < T$ and set $\theta = t_d^i - t$, $\theta_d^0 = T - t_d^1$, and for $j = 1 \dots i-1$, $\theta_d^j = t_d^j - t_d^{j+1}$, then the value at time t when the spot price of the stock is equal to x of the American Put option with strike K and maturity T is given by $u_i(\theta, x)$.*

With these notations, at time $t = t_d^i$, if the spot price of the stock is x, the price of the put option is $u_{i-1}(\theta_d^{i-1}, x)$. When $D_i(x)$ is positive, it differs from $u_i(0,x) = u_{i-1}(\theta_d^{i-1}, x - D_i(x))$. The next Lemma follows from Lemma 1.3 [JV11].

Lemma 3.2.2 *For each $\theta \geq 0$, the mapping $x \mapsto u_i(\theta, x)$ is non-increasing and $x \mapsto x + u_i(\theta, x)$ is non-decreasing.*

Like in Lemma 1.3 [JV11], one easily deduces the existence of the exercise boundary.

Corollary 3.2.3 (Exercise boundary) *For $i \in \{1,\dots,I\}$ and $\theta \geq 0$, it exists $c_i(\theta) \in [0, K)$ such that : $u_i(\theta, x) > (K-x)^+ \Leftrightarrow x > c_i(\theta)$*

By Proposition 3.2.1, the exercise boundary of the American Put option in our model with discrete dividends is

$$t \in [0,T) \mapsto \sum_{i=0}^{I} c_i(t_d^i - t)\mathbf{1}_{\{t_d^{i+1} \leq t < t_d^i\}} \text{ with convention } t_d^0 = T.$$

With a slight abuse of terminology, we also call exercise boundaries the functions c_i. Notice that because the argument of c_i is the time to the dividend date t_d^i, left-continuity of the c_i implies right-continuity of the true exercise boundary and that right-continuity of the c_i implies left-continuity of the true boundary on $[0, t_d^I) \cup (t_d^I, t_d^{I-1}) \cup \cdots \cup (t_d^1, T)$ with existence of left-hand limits at the dividend dates.

According to Lemma 1.4 [JV11], one has

Proposition 3.2.4 (Regularity result) *The value function $(\theta, x) \mapsto u_i(\theta, x)$ is continuous on $\mathbb{R}_+ \times \mathbb{R}_+$. On the continuation region defined as $\{(\theta, x) | \theta > 0,\ x > c_i(\theta)\}$, this function is $\mathcal{C}^{1,2}$ and satisfies :*

$$-\partial_\theta u_i(\theta, x) - r u_i(\theta, x) + r x \partial_x u_i(\theta, x) + \frac{\sigma^2}{2} x^2 \partial_{xx} u_i(\theta, x) = 0.$$

Moreover, the left-hand derivative $\partial_{xx-} u_i(\theta, x)$ of $\partial_x u_i(\theta, \bullet)$ is well defined and equal to 0 in the exercise region $\{(\theta, x) | \theta > 0,\ 0 \leq x \leq c_i(\theta)\}$.

The upper-semi continuity of $c_i(\bullet)$ is a consequence of the continuity of u_i.

Corollary 3.2.5 *For any $\theta \geq 0$, $\limsup_{\theta' \to \theta} c_i(\theta') \leq c_i(\theta)$.*

Remark 3.2.6 *Since the dividend function D_i is non-negative, $u_i(\theta, x) \geq u_{i-1}(\theta + \theta_d^{i-1}, x)$ and therefore $u_i(\theta, x) \geq u_0\left(\theta + \sum_{j=1}^{i} \theta_d^{i-1}, x\right)$. We deduce that $c_i(\theta) \leq K \bar{c}\left(\theta + \sum_{j=1}^{i} \theta_d^{i-1}\right)$. In particular, if $r = 0$, $\bar{c}(t) = 0$ for $t > 0$, so that $c_i \equiv 0$ for $i \in \{1, \ldots, I\}$.*

3.3 Smooth-fit property

In this section, we are going to prove the smooth-fit property. See [PS06, p.149] for a discussion of this property in optimal stopping problems and [Jac93, Prop.2.8], [PS06, p.375-395] or [KS91, p.73-79] for the specific case of the American Put option in the Black-Scholes model.

Proposition 3.3.1 (Smooth-fit) *For all $\theta > 0$, $u_i(\theta, \bullet)$ is $\mathcal{C}^1$.*

The proof is based on the viscosity super-solution property of u_i and estimations of the time derivative of this function stated in the two next Lemmas.

Lemma 3.3.2 *$(\theta, x) \mapsto u(\theta, x)$ is a viscosity supersolution of*
$\min(\partial_\theta u_i(\theta, x) - \mathcal{A} u_i(\theta, \bullet)(x), u_i(\theta, x) - (K - x)^+) = 0$ with $u_i(0, x) = u_{i-1}(\theta_d^{i-1}, \rho_i(x))$

Proof. It comes from the definition of u_i that $u_i(\theta, x) \geq (K - x)^+$.
Let $\phi(t, x)$ be a test function such that $0 = (u_i - \phi)(\theta, x) = \min_V (u_i - \phi)$ where $V = (\theta - \eta, \theta] \times (x - \eta, x + \eta)$ for a certain $\eta > 0$. Let τ be the first exit time of $\bar{S}^x$ outside the ball centered at x with radius η and let $0 < \epsilon < \eta$. Because of the minimum property of (θ, x), one has

$$\mathbb{E}\left[e^{-r(\tau \wedge \epsilon)}(u_i(\theta - (\tau \wedge \epsilon), \bar{S}^x_{\tau \wedge \epsilon}) - \phi(\theta - (\tau \wedge \epsilon), \bar{S}^x_{\tau \wedge \epsilon}))\right] \geq u_i(\theta, x) - \phi(\theta, x).$$

Applying Itô formula to $e^{-rt}\phi(\theta - t, \bar{S}^x_t)$ between $t = 0$ and $\tau \wedge \epsilon$, we deduce that

$$\mathbb{E}\left[\int_0^{\tau\wedge\epsilon} e^{-rt}(\partial_\theta\phi(\theta-t,\bar{S}^x_t)-\mathcal{A}\phi(\theta-t,\bullet)(\bar{S}^x_t))dt\right]\geq\mathbb{E}\left[\left(u_i(\theta,x)-e^{-r(\tau\wedge\epsilon)}u_i(\theta-(\tau\wedge\epsilon),\bar{S}^x_{\tau\wedge\epsilon})\right)\right].$$

Since, by the dynamic programming principle, for any stopping time $\eta\leq\theta$, one has $u_i(\theta,x)\geq\mathbb{E}\left[e^{-r\eta}u_i(\theta-\eta,\bar{S}^x_\eta)\right]$, the right-hand-side is non-negative. We deduce that

$$\mathbb{E}\left[\frac{1}{\epsilon}\int_0^{\tau\wedge\epsilon} e^{-rt}(\partial_\theta\phi(\theta-t,\bar{S}^x_t)-\mathcal{A}\phi(\theta-t,\bullet)(\bar{S}^x_t))dt\right]\geq 0.$$

By sending ϵ to zero, we obtain the supersolution inequality from Lebesgue's theorem :

$$\partial_\theta\phi(\theta,x)-\mathcal{A}\phi(\theta,\bullet)(x)\geq 0.$$

□

Lemma 3.3.3 *For any $i\geq 0$, $\theta>0$ and $x\geq 0$ one has*

$$\limsup_{\theta'\to\theta}\left|\frac{u_i(\theta',x)-u_i(\theta,x)}{\theta'-\theta}\right|\leq r\left(K+x\right)+x\left(r\left(2\mathcal{N}\left(\frac{2r}{\sigma}\sqrt{\theta}\right)-1\right)+\sigma\frac{e^{-2\frac{r^2}{\sigma^2}\theta}}{\sqrt{2\pi\theta}}\right),$$
$$|\partial_{xx-}u_i(\theta,x)|\leq\mathbf{1}_{\{x\geq c_i(\theta)\}}\frac{2}{\sigma^2c_i^2(\theta)}\left(2rK+\left(3r+\frac{\sigma}{\sqrt{2\pi\theta}}\right)c_i(\theta)\right).$$

Moreover $\partial_x u_i(\theta,x)$ admits a right-hand limit at $c_i(\theta)$ denoted by $\partial_x u_i(\theta,c_i(\theta)^+)$ and $\partial_x u_i(\theta,c_i(\theta)^+)\in[-1,0]$.

The proof of these estimates, which relies on the scaling property of the Brownian motion and Lemma 3.2.2, is postponed in Appendix. We are now able to prove Proposition 3.3.1. **Proof.** Let $c=c_i(\theta)$. By Lemma 3.3.3, the limit $\partial_x u_i(\theta,c+)=\lim_{y\downarrow c}\partial_x u_i(\theta,y)$ exists.
We adapt a viscosity solution argument given in [Pha09] : supposing that $\partial_x u(\theta,c+)>-1$, we are going to obtain a contradiction. For $\epsilon>0$, let $\phi_\epsilon(x)=(K-c)^++\alpha(x-c)+\frac{1}{2\epsilon}(x-c)^2$ where $-1=\partial_x u_i(\theta,c-)<\alpha<\partial_x u_i(\theta,c+)$. Since $c<K$, it exists an open interval $(x_\epsilon,y_\epsilon)\subset[0,K]$ containing c such that $\min_{x\in(x_\epsilon,y_\epsilon)}(u_i(\theta,x)-\phi_\epsilon(x))=u_i(\theta,c)-\phi_\epsilon(c)=0$.
We set

$$\beta=2(3r+\tfrac{\sigma}{\sqrt{\pi\theta}})K\ \text{ and }\ \phi(\theta-t,x)=\phi_\epsilon(x)-\beta t.$$

By Lemma 3.3.3, for any $(t,x)\in[0,\frac{\theta}{2}]\times[0,K]$, one has $u_i(\theta-t,x)-u_i(\theta,x)\geq-\dfrac{\beta}{2}t$. Therefore $0=(u_i-\phi)(\theta,c)=\min_{(t,x)\in(\frac{\theta}{2},\theta]\times(x_\epsilon,y_\epsilon)}(u_i-\phi)(t,x)$. By the supersolution property of u_i stated in Lemma 3.3.2, we deduce that

$$0\leq\partial_\theta\phi(\theta,c)-\mathcal{A}\phi(\theta,\bullet)(c)=\beta+r(K-c)-rc\alpha-\frac{\sigma^2c^2}{2\epsilon}.$$

By sending ϵ to zero, we get the desired contradiction. □

3.4 Continuity of the exercise boundary

Proposition 3.4.1 *Under (A), for any $i \in \{0, \dots, I\}$, the function $\theta \mapsto c_i(\theta)$ is continuous on $[0, +\infty)$.*

The right continuity will be proved in subsection 3.4.1 whereas the left continuity will be proved in subsection 3.4.2.

Remark 3.4.2 *In particular, we deduce from this result the behaviour of the exercise boundary at the dividend time.*

Since $c_i(0) = \sup\left\{x \geq 0 | u_{i-1}(\theta_d^{i-1}, x - D_i(x)) = K - x\right\}$ and for $y \in [0, c_{i-1}(\theta_d^{i-1}))$, $u_{i-1}(\theta_d^{i-1}, y) = K - y$, one has $c_i(0) = c_{i-1}(\theta_d^{i-1}) \wedge \inf\{x \geq 0 | D_i(x) > 0\}$ and

Corollary 3.4.3 *Under (A), for any $i \in \{1, \dots, I\}$, $\lim_{t \to 0_+} c_i(t) = c_{i-1}(\theta_d^{i-1}) \wedge \inf\{x \geq 0 | D_i(x) > 0\}$.*

As $c_i(0) = 0$ when $\forall x > 0$, $D_i(x) > 0$, this result generalizes Lemma 2.1 [JV11].

3.4.1 Right continuity

The right continuity of the exercise boundary is based on a comparaison result with the exercise boundary $\bar{c}$ of the classical American Put option with strike 1 in the Black-Scholes model without dividends.

Lemma 3.4.4 *For $\theta \geq 0$ and $t \geq 0$, one has : $c_i(\theta + t) \geq \left(K\left(1 - e^{-rt}\right) + c_i(\theta)e^{-rt}\right)\bar{c}(t)$*

Proof. Let $\tau = \tilde{\tau} \wedge t$ where $\tilde{\tau}$ is an optimal stopping time for $u_i(\theta + t, x)$. By the dynamic programming principle, one has

$$u_i(\theta + t, x) = \mathbb{E}\left[e^{-r\tau}(K - \bar{S}_\tau^x)^+ \mathbf{1}_{\{\tau < t\}} + \mathbf{1}_{\{\tau = t\}} e^{-rt} u_i(\theta, \bar{S}_t^x)\right].$$

Since $x \mapsto u_i(\theta, x)$ is non-increasing and using the fact for any $0 \leq \alpha \leq K$, $(K - x)^+ \leq (K - (\alpha \wedge x))^+ = (K - \alpha) + (\alpha - x)^+$, one deduces

$$
\begin{aligned}
u_i(\theta+t,x) &\leq \mathbb{E}\left[e^{-r\tau}(K-\bar{S}^x_\tau)^+\mathbf{1}_{\{\tau<t\}}+\mathbf{1}_{\{\tau=t\}}e^{-rt}\left(K-c_i(\theta)\wedge\bar{S}^x_t\right)^+\right]\\
&\leq \mathbb{E}\left[\begin{array}{l} e^{-r\tau}\left(K-\left\{c_i(\theta)+(K-c_i(\theta))\left(1-e^{-r(t-\tau)}\right)\right\}\wedge\bar{S}^x_\tau\right)^+\mathbf{1}_{\{\tau<t\}}\\ +\mathbf{1}_{\{\tau=t\}}e^{-rt}\left(K-c_i(\theta)\wedge\bar{S}^x_t\right)^+\end{array}\right]\\
&\leq \mathbb{E}\left[e^{-r\tau}\left(K-\left\{c_i(\theta)+(K-c_i(\theta))\left(1-e^{-r(t-\tau)}\right)\right\}\wedge\bar{S}^x_\tau\right)^+\right]\\
&\leq \mathbb{E}\left[e^{-r\tau}\left(K-\left\{c_i(\theta)+(K-c_i(\theta))\left(1-e^{-r(t-\tau)}\right)\right\}\right)\right]\\
&\qquad +\mathbb{E}\left[e^{-r\tau}\left(c_i(\theta)+(K-c_i(\theta))\left(1-e^{-r(t-\tau)}\right)-\bar{S}^x_\tau\right)^+\right]\\
&\leq (K-c_i(\theta))e^{-rt}+\mathbb{E}\left[e^{-r\tau}\left(K\left(1-e^{-rt}\right)+c_i(\theta)e^{-rt}-\bar{S}^x_\tau\right)^+\right]
\end{aligned}
$$

where we used $(K-c_i(\theta))(1-e^{-r(t-\tau)})\leq(K-c_i(\theta))(1-e^{-rt})$ for the last inequality.

Since τ is a stopping-time not greater then t, for $x\leq\left(K\left(1-e^{-rt}\right)+c_i(\theta)e^{-rt}\right)\bar{c}(t)$, the second term of the right-hand side is not greater than $(K\left(1-e^{-rt}\right)+c_i(\theta)e^{-rt}-x)$. Therefore, one has $u_i(\theta+t,x)\leq(K-x)^+$ and $c_i(\theta+t)\geq x$. □

Corollary 3.4.5 *The function $\theta\mapsto c_i(\theta)$ is right continuous.*

Proof. Because $\lim_{t\to0}\bar{c}(t)=1$ (cf [KS91] p.71-80), Lemma 3.4.4 implies that $\liminf_{\theta'\downarrow\theta}c_i(\theta')\geq c_i(\theta)$. We conclude with the upper-semicontinuity property stated in Corollary 3.2.5. □

We recall (cf [KS91]) that $\bar{c}(\infty)\stackrel{\text{def}}{=}\lim_{\theta\to+\infty}\bar{c}(\theta)$ exists and is equal to $\frac{2r}{2r+\sigma^2}$.

Corollary 3.4.6 *One has $\lim_{\theta\to+\infty}c_i(\theta)=K\bar{c}(\infty)$. Moreover, when $r>0$, $\forall\theta>0$, $c_i(\theta)>0$.*

Proof. If $r=0$ then by Remark 3.2.6 the statement clearly holds.
Let us now assume that $r>0$. Since $u_i(t,x)\geq u_0(t,x)$, we have $c_i(t)\leq K\bar{c}(t)$. Writing Lemma 3.4.4 for $\theta=0$, we deduce that

$$\forall t\geq0,\ -(K-c_i(0))e^{-rt}\bar{c}(t)\leq c_i(t)-K\bar{c}(t)\leq0.$$

We obtain the first statement by taking the limit $t\to\infty$ in this inequality.
For $\theta=0$, Lemma 3.4.4 also implies $c_i(t)\geq K(1-e^{-rt})\bar{c}(t)$. Since $\bar{c}$ is non-increasing with positive limit at infinity, we deduce that $c_i(t)>0$ as soon as $t>0$. □

3.4.2 Left continuity

The left continuity is based on the characterization of the continuation region in terms of the spatial derivative of u_i stated in the next proposition.

Proposition 3.4.7 *Under (A), the property*

$$(P_i) : \textit{For any } \theta > 0 \textit{ and } x \geq 0 \textit{ one has } x > c_i(\theta) \Longleftrightarrow 1 + \partial_x u_i(\theta, x) > 0$$

holds for any $i \in \{0, \cdots, I\}$.

The proof of Proposition 3.4.7 will be done by induction on i. The main tools to deduce the induction hypothesis at rank i from the one at rank $i-1$ are in the following Lemmas, the proofs of which are postponed to the Appendix.

Lemma 3.4.8 *Let $\theta > 0$, $x > c_i(\theta)$ and τ denote the smallest optimal stopping time for $u_i(\theta, x)$. Then $y \mapsto \mathbb{P}\left(\tau = \theta | \bar{S}^x_\theta = y\right)$ is non-decreasing and is positive on $(K, +\infty)$.*

The function $u_i(0, x)$ being Lipschitz continuous by Lemma 3.2.2, it is absolutely continuous and therefore dx a.e. differentiable. We denote by $\partial_x u_i(0, x)$ its a.e. derivative. For $\theta > 0$ and $x > 0$, since $\bar{S}^x_\theta$ admits a density with respect to the Lebesgue measure under $\mathbb{P}$, the random variable $\partial_x u_i(0, \bar{S}^x_\theta)$ is a.s defined under $\mathbb{P}$ and therefore under $\mathbb{Q}$.

Lemma 3.4.9 *Let $\theta > 0$, $x \geq 0$ and τ be an optimal stopping time for $u_i(\theta, x)$. Then one has*

$$1 + \partial_x u_i(\theta, x) \geq \mathbb{E}^{\mathbb{Q}}\left[\mathbf{1}_{\{\tau=\theta\}}\left(1 + \partial_x u_i(0, \bar{S}^x_\theta)\right)\right].$$

Moreover, $\overline{\tau} \stackrel{\text{def}}{=} \lim_{\epsilon \to 0_+} \inf\left\{t \geq 0 | \bar{S}^{x+\epsilon}_t \leq c_i(\theta - t)\right\}$ is an optimal stopping time and satisfies

$$1 + \partial_x u_i(\theta, x) = \mathbb{E}^{\mathbb{Q}}\left[\mathbf{1}_{\{\overline{\tau}=\theta\}}\left(1 + \partial_x u_i(0, \bar{S}^x_\theta)\right)\right].$$

We are now proving Proposition 3.4.7. **Proof.** First, for $i = 0$, due to [KS91], $x \mapsto u_i(\theta, x)$ is convex and so (P_0) is true.
Let us suppose that (P_{i-1}) holds for $i \in \{1, \cdots, I-1\}$.
By (A), $\kappa_i \stackrel{\text{def}}{=} \sup\left\{x \geq 0 | x - D_i(x) \leq c_{i-1}(\theta^{i-1}_d)\right\}$ is such that

$$\forall x \geq 0,\ x - D_i(x) \leq c_{i-1}(\theta^{i-1}_d) \Leftrightarrow x \leq \kappa_i.$$

Moreover, D_i is differentiable dx a.e. and equal to the integral of its a.e. derivative which takes its values in $[0, 1]$. We denote this a.e. derivative by D'_i. Since $u_i(0, x) = u_{i-1}(\theta^{i-1}_d, x - D_i(x))$ where $u_{i-1}(\theta^{i-1}_d, x)$ is C^1 by Proposition 3.3.1, one easily checks that

$$dx \text{ a.e., } \partial_x u_i(0, x) = (1 - D'_i(x))\partial_y u_{i-1}(\theta^{i-1}_d, y)|_{y=x-D_i(x)} \tag{3.1}$$

where the second term of the right-hand-side belongs to $[-1, 0]$ by Lemma 3.2.2. There are two possibilities :

- either $\kappa_i < \infty$ and then for $x > \kappa_i$, $1 + \partial_y u_{i-1}(\theta_d^{i-1}, y)|_{y=x-D_i(x)} > 0$ by (P_{i-1}) so that $1 + \partial_x u_i(0,x) > 0$ a.e. by Equation 3.1,
- or $\kappa_i = +\infty$ and then $D_i(x) = \int_0^x D_i'(y)dy \sim x$ as $x \to \infty$. Therefore there exists a borel set $\mathbf{C} \subset (K, +\infty)$ with infinite Lebesgue measure, on which D_i' takes values in $\left[\frac{1}{2}, 1\right]$. By Equation 3.1, for almost every $x \in \mathbf{C}$, $1 + \partial_x u_i(0,x) \geq \frac{1}{2}$.

So there exists of a borel set $A \subset (K, +\infty)$ which is non neglictible for the Lebesgue measure and such that for every $x \in A$, $1 + \partial_x u_i(0,x) > 0$.
Using the first statement of Lemma 3.4.9 then $\frac{d\mathbb{Q}}{d\mathbb{P}}|_{\mathcal{F}_\theta} = \frac{e^{-r\theta}\bar{S}^x_\theta}{x}$, one obtains

$$\begin{aligned}
1 + \partial_x u_i(\theta, x) \geq \; & \mathbb{E}^{\mathbb{Q}}\left[\mathbf{1}_{\{\tau=\theta\}}\left(1 + \partial_x u_i(0, \bar{S}^x_\theta)\right)\right] \\
= \; & e^{-r\theta}\int_0^{+\infty} \frac{y}{x}\left(1 + \partial_x u_i(0,y)\right)\mathbb{P}\left(\tau = \theta | \bar{S}^x_\theta = y\right)\mathrm{p}(\theta, x, y)dy \\
\geq \; & e^{-r\theta}\int_A \frac{y}{x}\left(1 + \partial_x u_i(0,y)\right)\mathbb{P}\left(\tau = \theta | \bar{S}^x_\theta = y\right)\mathrm{p}(\theta, x, y)dy.
\end{aligned}$$

By Lemma 3.4.8, the last quantity is positive and the assertion is proved. □

Proposition 3.4.10 *$\theta \mapsto c_i(\theta)$ is left continuous.*

Proof. When $r = 0$, by Remark 3.2.6, the statement holds. Let us assume that $r > 0$. By Corollary 3.2.5, we just need to prove that it does not exist $\theta > 0$ such that $\liminf_{t\to 0_+} c_i(\theta - t) < c_i(\theta)$.
Let us suppose that it exists such a $\theta > 0$ and obtain a contradiction. Let $c_- \stackrel{\text{def}}{=} \liminf_{t\to 0_+} c_i(\theta - t)$ and $(t_n)_n$ be a decreasing sequence in $(0, \theta)$ tending to zero and such that $c_i(\theta - t_n)$ tend to c_-. Then, by Lemma 3.4.4 written with $(s - t_n, \theta - s)$ replacing (t, θ), we obtain that for $s \in (t_n, \theta)$, $c_i(\theta - s) \leq c_i(\theta - t_n)\dfrac{e^{r(s-t_n)}}{\bar{c}(s-t_n)}$. So $\lim_{t\to 0_+} c_i(\theta - t) = c_-$. Then it exists $\eta \in (0, c_i(\theta))$, $\delta_0 \in (0, \theta/2)$, such that $\forall t \in (0, 2\delta_0)$ $c_i(\theta - t) < c_i(\theta) - \eta$. Let $x < y$ be such that $c_i(\theta) - \eta < x < y \leq c_i(\theta)$. One has

$$y - x + u_i(\theta, y) - u_i(\theta, x) = 0. \tag{3.2}$$

Let us define $\tau = \inf\left\{t \geq 0 \;\middle|\; t + \left|\bar{S}^1_t - 1\right| \geq \delta_0 \wedge \frac{x - c_i(\theta)+\eta}{x}\right\}$. For $\theta' \in (\theta, \theta - \delta_0)$ and $z \geq x$, one has $\forall t \in [0, \tau]$, $\bar{S}^z_t \geq \bar{S}^x_t \geq c_i(\theta) - \eta > c_i(\theta' - t)$ and by Proposition 3.2.4, $u_i(\theta', z) = \mathbb{E}\left[e^{-r\tau}u_i(\theta' - \tau, \bar{S}^z_\tau)\right]$. Since u_i is continuous and bounded by K, letting θ' tend to θ, we get by dominated convergence $u_i(\theta, z) = \mathbb{E}\left[e^{-r\tau}u_i(\theta - \tau, \bar{S}^z_\tau)\right]$. We deduce

$$\begin{aligned}
y - x + u_i(\theta, y) - u_i(\theta, x) &= \mathbb{E}\left[e^{-r\tau}\left(\bar{S}^y_\tau - \bar{S}^x_\tau + u_i(\theta - \tau, \bar{S}^y_\tau) - u_i(\theta - \tau, \bar{S}^x_\tau)\right)\right] \\
&= \mathbb{E}^{\mathbb{Q}}\left[\int_x^y \left(1 + \partial_x u_i(\theta - \tau, \bar{S}^z_\tau)\right)dz\right].
\end{aligned}$$

But since $\mathbb{Q}\left(\tau > 0 \text{ and } \forall z \geq x,\ \bar{S}^z_\tau > c_i(\theta - \tau)\right) = 1$, the right-hand side is positive by Proposition 3.4.7, which contradicts Equation 3.2. □

On Figure 3.2, we represent two different exercise boundaries computed through a binomial tree method following [VN06]. In both cases, $c_1(0) = \kappa_1 = 20$. In case (a), the boundary appears to be smooth whereas in case (b), it seems to be merely continuous (at time 0.04, even continuity is not so clear from the figure).

(a) Maturity is 2 with one dividend time at 1.7; $D_1(x) = \frac{1}{5}\left((x-20)^+ - (x-30)^+\right)$

(b) Maturity is 0.1 with one dividend time at 0.05; $D_1(x) = \min\left(\frac{9}{8}, \frac{2}{9}\left((x-20)^+\right)^2\right)$

Fig. 3.2. Exercise boundaries of an American Put option with different maturities for different dividend functions. Strike is 100, diffusion parameters are $r = 0.04$ and $\sigma = 0.3$.

3.5 Local behaviour of the exercise boundary near the dividend dates

In this section, we are going to show how the behaviour of the exercise boundary is driven by the shape of the function $u_i(0,.)$ when $i \in \{1,\dots,I\}$.
We recall that $c_i(0) = \min\left(c_{i-1}(\theta_d^{i-1}), \inf\{x \geq 0 | D_i(x) > 0\}\right)$. Applying Lemma 3.4.4 for $\theta = 0$ and $t = \theta_d^{i-1}$, one obtains

$$c_{i-1}(\theta_d^{i-1}) \geq \left(K(1 - e^{-r\theta_d^{i-1}}) + e^{-r\theta_d^{i-1}} c_{i-1}(0)\right)\bar{c}(\theta_d^{i-1}) \geq \frac{2rK}{2r+\sigma^2}(1 - e^{-r\theta_d^{i-1}}) \tag{3.3}$$

We are able to precise the local behaviour of the exercise boundary near the dividend dates only when $c_i(0) < c_{i-1}(\theta_d^{i-1})$ which is satisfied as soon as $\inf\{x \geq 0 | D_i(x) > 0\} < \frac{2rK}{2r+\sigma^2}(1 - e^{-r\theta_d^{i-1}})$. On Figure 3.3 are represented two different exercise boundaries computed through a binomial tree method following [VN06]. Notice that in each case, a dividend is paid if the stock price is over 50. On the left (resp. right) one, $c_1(.)$ seems to be locally increasing (resp. decreasing) on $[0,\epsilon)$ for ϵ small enough. In Propositions 3.5.3 and 3.5.6, we give sufficient conditions on the dividend functions for these local monotonicity properties to hold.

(a) $D_1(x) = \min\left(\frac{1125}{32}, \frac{8}{1125}\left((x-50)^+\right)^2\right)$ (b) $D_1(x) = 0.05(x-50)^+$

Fig. 3.3. Exercise boundaries of an American Put option of maturity 4 with one dividend time at 3.5 for different dividend functions. Strike is 100, diffusion parameters are $r = 0.04$ and $\sigma = 0.3$.

3.5.1 Equivalent of the exercise boundary for dividend functions with positive slope at $c_i(0)_+$

Proposition 3.5.1 *If $c_i(0) > 0$ and $\liminf_{x\to c_i(0)_+} \frac{D_i(x)}{x-c_i(0)} > 0$, then we have that $c_i(\theta) - c_i(0) \sim_{\theta\to 0_+} -\sigma c_i(0)\sqrt{\theta\,|\ln\theta|}$.*

By Remark 3.2.6 and Equation 3.3, a necessary and sufficient condition for the positivity of $c_i(0)$ is positivity of both r and $\inf\{x \geq 0 | D_i(x) > 0\}$. Notice that the second hypothesis implies that $c_i(0) = \inf\{x \geq 0 | D_i(x) > 0\}$ and therefore that $\inf\{x \geq 0 | D_i(x) > 0\} \leq c_{i-1}(\theta_d^{i-1})$ with possible equality. In order to prove Proposition 3.5.1, we need the following Lemma, the proof of which is postponed in Appendix.

Lemma 3.5.2 *Suppose that $c_i(0) > 0$ and that it exists $\alpha > 0$, $\beta \in [1,2)$ and an open set $V \subset \mathbb{R}_+^\star$ containing $c_i(0)$ such that :*

$$\forall x \in V,\ u_i(0,x) - (K-x)^+ \geq \alpha \left|(x - c_i(0))^+\right|^\beta. \tag{3.4}$$

Then $\forall \delta > 1,\ \exists \Theta_\delta > 0,\ \forall \theta \in [0, \Theta_\delta],\ c_i(\theta) \leq c_i(0) \exp\left\{-\sigma\sqrt{\theta\left((2-\beta)\,|\ln\theta| - (\beta+\delta)\ln|\ln\theta|\right)}\right\}.$

In particular, when $D_i(x) = \alpha(x-\beta)^+ \wedge \gamma$ with $\alpha \in (0,1)$, $\beta \in (0, c_{i-1}(\theta_d^{i-1})]$ and $\gamma > 0$, in a neighborhood of 0, the exercise boundary c_i is under a decreasing function coinciding with $c_i(0)$ at 0.

We are now able to prove Proposition 3.5.1. **Proof.** Since $c_i(0) \leq c_{i-1}(\theta_d^{i-1}) < K$ and for $x \in [0,K]$, $u_{i-1}(0,x) \geq K - x + D_i(x)$, the positivity of $\liminf_{x\to c_i(0)_+} \frac{D_i(x)}{x-c_i(0)}$ implies that the second hypothesis of Lemma 3.5.2 is satisfied with $\beta = 1$. Hence, for θ small enough, $c_i(\theta) \leq c_i(0)e^{-\sigma\sqrt{\theta(|\ln\theta| - 3\ln|\ln\theta|)}}$. By Lemma 3.4.4, we know that $c_i(\theta) \geq c_i(0)\bar{c}(\theta) + (1 - e^{-r\theta})\,(K - c_i(0))\,\bar{c}(\theta)$, where, according to [Lam95], $\bar{c}(\theta) - 1 \sim_{\theta\downarrow 0} -\sigma\sqrt{\theta\,|\ln\theta|}$. Since $\sqrt{\theta\,(|\ln\theta| - 3\ln|\ln\theta|)} \sim_{\theta\downarrow 0} \sqrt{\theta\,|\ln\theta|}$, we easily conclude. □

3.5.2 Monotonicity of the value function

The monotonicity of the value function around the i-th dividend time is closely related to the sign, on a right-hand neighborhood of $c_i(0)$, of the Black-Scholes operator applied to $u_i(0,.) = u_{i-1}(\theta_d^{i-1}, \rho_i(.))$ where $\rho_i(x) = x - D_i(x)$. In the previous sections, the derivative of D_i was thought in the sense of distributions. From now on, we assume that D_i is the difference of two convex functions in order to apply the Itô-Tanaka formula. So the derivative of D_i (resp. ρ_i) is considered as the left-hand derivative.

Exercise boundary locally non-decreasing

To obtain this property, we need negativity of the Black-Scholes operator applied to $u_i(0,.)$ in a right-hand neighborhood of $c_i(0)$.

Proposition 3.5.3 *Assume that* $\inf\{x \geq 0 | D_i(x) > 0\} < c_{i-1}(\theta_d^{i-1})$, *that D_i is the difference of two convex functions, and that the positive part of the Jordan-Hahn decomposition of the measure D_i'' is absolutely continuous with respect to the Lebesgue measure. Assume moreover that, if g_i denotes the density of the absolutely continuous part of D_i'', it exists $\varepsilon \in (0, c_{i-1}(\theta_d^{i-1}) - c_i(0))$ and $C_1 \in [0,+\infty)$ such that*

$$\begin{aligned}
&\forall x \leq c_i(0) + \varepsilon,\ -rD_i(x) + rxD_i'(x) + \frac{\sigma^2 x^2}{2} g_i(x) \leq rK - \varepsilon \\
&\forall x > c_i(0) + \varepsilon,\ g_i(x) \leq C_1 x^{C_1}.
\end{aligned}$$

Then it exists a neighborhood of $(0, c_i(0))$ in $\mathbb{R}_+ \times \mathbb{R}_+$ such that u_i is non-increasing w.r.t θ in this neighborhood. Moreover, the exercise boundary c_i is non-decreasing in a neighborhood of 0.

According to Equation 3.3, when $\inf\{x \geq 0 | D_i(x) > 0\} < \frac{2rK}{2r+\sigma^2}(1 - e^{-r\theta_d^{i-1}})$, then $\inf\{x \geq 0 | D_i(x) > 0\} < c_{i-1}(\theta_d^{i-1})$.

Remark 3.5.4 *This result is a generalization of Proposition 2.2 in [JV11] which states the same local monotonicity property of the value function at the first dividend date when $c_1(0) = 0$ and D_1 is a non-zero concave function satisfying assumption (A). Indeed concavity implies that $g_1(x) \leq 0$ and $D_1(x) - rxD_1'(x) \geq D_1(0)$ where $D_1(0) = 0$ by (A). When $r > 0$ and $c_i(0) = 0$, generalizing the proofs of Lemma 2.1 and Corollary 2.3 [JV11], one may check that $c_i(\theta) \leq rK\theta \limsup_{x\to 0^+} \frac{x}{D_i(x)} + o(\theta)$ as*

$\theta \to 0$ and that, under the assumptions of Proposition 3.5.3, if $\frac{x}{D_i(x)}$ admits a finite right-hand limit at $x = 0$, $c_i(\theta) \sim_{\theta\to 0^+} rK\theta \lim_{x\to 0^+} \frac{x}{D_i(x)}$.

When $r > 0$, for $\beta \in (0, c_{i-1}(\theta_d^{i-1}))$, $\eta \in (0, r)$ and $\alpha \in (0, \frac{\sigma^2\beta^2}{4(r-\eta)K}]$, the function $D_i(x) = \min\left(\alpha, \frac{(r-\eta)K}{\sigma^2\beta^2}\left((x-\beta)^+\right)^2\right)$ satisfies (A) and the assumptions of Proposition 3.5.3.

To prove the Proposition, we need the following Lemma, the proof of which is postponed in appendix.

Lemma 3.5.5 *For $t_1 \geq 0$, let $\tau_{t_1} = \inf\left\{w \geq 0 | \bar{S}_w^x \geq c_i(t_1 - w)\mathbf{1}_{\{w<t_1\}} + c_i(0)\mathbf{1}_{\{w \geq t_1\}}\right\}$ with the convention $\inf\emptyset = +\infty$.*

$$\forall p \geq 0,\ \forall \alpha > 0,\ \exists \eta > 0,\ \lim_{v\to 0+} \sup_{t_1 \leq \eta} \sup_{x \leq c_i(0)+\alpha} \frac{\mathbb{E}\left[\left(1 + \left(\bar{S}_v^x\right)^p\right)\mathbf{1}_{\{\tau_{t_1}\geq v,\ \bar{S}_v^x \geq c_i(0)+2\alpha\}}\right]}{\mathbb{P}(\tau_{t_1} \geq v)} = 0.$$

We are now able to prove Proposition 3.5.3. **Proof.** Let $0 \leq s < t$, $x > c_i(t)$ and τ be the smallest optimal stopping time for (t, x). Since $\tau \wedge s$ is a stopping time not greater than s, $u_i(s, x) \geq \mathbb{E}\left[e^{-r\tau}\left(K - \bar{S}_\tau^x\right)\mathbf{1}_{\{\tau<s\}} + e^{-rs}u_i(0, \bar{S}_s^x)\right]$. Using $(K - x)^+ \leq u_i(0, x)$, we deduce

$$u_i(t, x) - u_i(s, x) \leq \mathbb{E}\left[\mathbf{1}_{\{\tau\geq s\}}\left(e^{-r\tau}u_i(0, \bar{S}_\tau^x) - e^{-rs}u_i(0, \bar{S}_s^x)\right)\right].$$

By Lemma 3.6.1, on $\tau > s$,

$$\begin{aligned} e^{-r\tau}u_i(0, \bar{S}_\tau^x) - e^{-rs}u_i(0, \bar{S}_s^x) = &\int_s^\tau e^{-rv}\left\{\begin{array}{l} -ru_i(0, \bar{S}_v^x) + r\bar{S}_v^x \partial_x u_{i-1}(\theta_d^{i-1}, \rho_i(\bar{S}_v^x))\rho_i'(\bar{S}_v^x) \\ +\frac{\sigma^2}{2}\left(\bar{S}_v^x \rho_i'\left(\bar{S}_v^x\right)\right)^2 \partial_{xx}u_{i-1}(\theta_d^{i-1}, \rho_i(\bar{S}_v^x)) \end{array}\right\} dv \\ &+ \frac{1}{2}\int_s^\tau \int_{\mathbb{R}} e^{-rv}\partial_x u_{i-1}(\theta_d^{i-1}, \rho_i(a))\rho_i''(da)dL_v^a(\bar{S}^x) \\ &+ M_\tau - M_s \end{aligned} \tag{3.5}$$

where $M_t = \int_0^t \sigma e^{-rv}\bar{S}_v^x \partial_x u_{i-1}(\theta_d^{i-1}, \rho_i(\bar{S}_v^x))\rho_i'(\bar{S}_v^x)dB_v$. As $\mathbb{E}\left[\langle M\rangle_t\right] \leq \sigma^2 t x^2 e^{\sigma^2 t}$, M_t is a true martingale and

$$\mathbb{E}\left[\mathbf{1}_{\{\tau\geq s\}}(M_\tau - M_s)\right] = \mathbb{E}\left[\mathbf{1}_{\{\tau\geq s\}}(\mathbb{E}\left[M_\tau|\mathcal{F}_s\right] - M_s)\right] = 0. \tag{3.6}$$

The function $y \mapsto \partial_x u_{i-1}(\theta_d^{i-1}, \rho_i(y))$ belongs to $[-1, 0]$ by Lemma 3.2.2 and is equal to -1 on $[0, c_i(0) + \varepsilon]$ since then $\rho_i(y) \leq y \leq c_i(0) + \varepsilon < c_{i-1}(\theta_{i-1}^d)$. Since for any $a \geq 0$, $t \mapsto L_t^a$ is a non-decreasing process and $\rho_i'' = -D_i''$, using the growth assumption on g_i, we deduce that $\mathbb{P}$-almost surely

$$\begin{aligned} &\int_s^\tau \int_{\mathbb{R}} e^{-rv}\partial_x u_{i-1}(\theta_d^{i-1}, \rho_i(a))\rho_i''(da)dL_v^a(\bar{S}^x) \\ &\qquad \leq \int_s^\tau \int_{\mathbb{R}} e^{-rv}\left(\mathbf{1}_{\{a\leq c_i(0)+\varepsilon\}}g_i(a) + \mathbf{1}_{\{a>c_i(0)+\varepsilon\}}C_1 a^{C_1}\right)da dL_v^a(\bar{S}^x) \end{aligned}$$

Using Exercise 1.15 p.232 [RY91], we deduce that

$$\int_s^\tau \int_{\mathbb{R}} e^{-rv}\partial_x u_{i-1}(\theta_d^{i-1},\rho_i(a))\rho_i''(da)dL_v^a(\bar{S}^x)$$
$$\leq \int_s^\tau \sigma^2 e^{-rv}\left(\bar{S}_v^x\right)^2(\mathbf{1}_{\left\{\bar{S}_x^x\leq c_i(0)+\varepsilon\right\}}g_i(\bar{S}_v^x)+C_1\mathbf{1}_{\left\{\bar{S}_v^x>c_i(0)+\varepsilon\right\}}(\bar{S}_v^x)^{C_1})dv. \tag{3.7}$$

By Lemma 3.3.3 and since $c_i(0)+\varepsilon<c_{i-1}(\theta_d^{i-1})$, it exists a finite constant C_2 not depending on s and t such that

$$\int_s^\tau e^{-rv}\left(\bar{S}_v^x\rho_i'\left(\bar{S}_v^x\right)\right)^2\partial_{xx}u_{i-1}(\theta_d^{i-1},\rho_i(\bar{S}_v^x))dv\leq C_2\int_s^\tau e^{-rv}\left(\bar{S}_v^x\right)^2\mathbf{1}_{\left\{\bar{S}_v^x>c_i(0)+\varepsilon\right\}}dv. \tag{3.8}$$

For $y\leq c_i(0)+\varepsilon$, $u_{i-1}(\theta_d^{i-1},\rho_i(y))=K-\rho_i(y)$ and

$$-ru_i(0,y)+ry\partial_x u_{i-1}(\theta_d^{i-1},\rho_i(y))\rho_i'(y)=-rK-rD_i(y)+ryD_i'(y)$$

where D_i is equal to 0 on $[0,c_i(0)]$. Hence the assumptions ensure that

$$-ru_i(0,y)+ry\partial_x u_{i-1}(\theta_d^{i-1},\rho_i(y))\rho_i'(y)+\frac{\sigma^2y^2}{2}g_i(y)\leq\begin{cases}-rK \text{ if } y\leq c_i(0)\\ -\varepsilon \text{ if } y\in(c_i(0),c_i(0)+\varepsilon]\end{cases} \tag{3.9}$$

When $y>c_i(0)+\varepsilon$, since $\partial_x u_{i-1}\leq 0$ and $\rho_i'\geq 0$, $-ru_i(0,y)+ry\partial_x u_{i-1}(\theta_d^{i-1},\rho_i(y))\rho_i'(y)$ is non-positive.

Taking expectations in Equation 3.5 and using Equations 3.6, 3.7, 3.8, 3.9, we deduce that it exists a constant $M>0$ such that

$$u_i(t,x)-u_i(s,x)\leq\int_s^t e^{-rv}\mathbb{P}(\tau\geq v)\left\{\begin{array}{l}-(rK\wedge\varepsilon)\\ +M\frac{\mathbb{E}\left[\mathbf{1}_{\left\{\tau\geq v,\bar{S}_v^x>c_i(0)+\varepsilon\right\}}\left(1+\left(\bar{S}_v^x\right)^{2+C_1}\right)\right]}{\mathbb{P}(\tau\geq v)}\end{array}\right\}dv \tag{3.10}$$

Applying Lemma 3.5.5 (with $p=2+C_1$, $t_1=t$ and $\alpha=\frac{\varepsilon}{2}$), we obtain that for t small enough, uniformly in $x\leq c_i(0)+\frac{\varepsilon}{2}$, the right-hand-side of Equation 3.10 is non-positive.
With Proposition 3.4.1, we deduce the existence of $\eta>0$ such that $\sup_{w\in[0,\eta]}c_i(w)\leq c_i(0)+\frac{\epsilon}{2}$ and

$$\forall 0\leq s<t<\eta,\ \forall x\in(c_i(t),c_i(0)+\frac{\epsilon}{2}],\ u_i(t,x)\leq u_i(s,x).$$

This inequality is still true for $x\leq c_i(t)$ since then $u_i(t,x)=(K-x)^+\leq u_i(s,x)$. For $0\leq s<t<\eta$, we conclude that $u_i(t,c_i(s))\leq u_i(s,c_i(s))=K-c_i(s)$, which implies that $c_i(s)\leq c_i(t)$. □

Exercise boundary locally non-increasing

To obtain this property, we need positivity of the Black-Scholes operator applied to $u_i(0,.)$ in a right-hand neighborhood of $c_i(0)$.

Proposition 3.5.6 *Assume that* $0 < \inf\{x \geq 0 | D_i(x) > 0\} < c_{i-1}(\theta_d^{i-1})$, *that* D_i *is the difference of two convex functions, and that the negative part of the Jordan-Hahn decomposition of the measure* D_i'' *is absolutely continuous with respect to the Lebesgue measure.*
Assume moreover that, if g_i *denotes the density of the absolutely continuous part of the measure* D_i'', *it exists* $\varepsilon \in (0, c_{i-1}(\theta_d^{i-1}) - c_i(0))$ *and* $C_1 \in [0, +\infty)$ *such that*

$$\text{on } (c_i(0), c_i(0)+\varepsilon],\ D_i \text{ is } C^2 \text{ and such that} - rD_i(x) + rxD_i'(x) + \frac{\sigma^2 x^2}{2} g_i(x) \geq rK + \varepsilon,$$
$$\forall x > c_i(0)+\varepsilon,\ g_i(x) \leq -C_1 x^{C_1}.$$

Then it exists a neighborhood of $(0, c_i(0))$ *in* $\mathbb{R}_+ \times \mathbb{R}_+$ *such that* u_i *is non-decreasing w.r.t* θ *in this neighborhood. Moreover the exercise boundary* c_i *is non-increasing in a neighborhood of* 0.

Remark 3.5.7 *When* $c_i(0) = 0$, *there is no non-negative function* D_i *satisfying the differential inequality on some interval* $(c_i(0), c_i(0)+\varepsilon)$. *That is why we suppose* $\inf\{x \geq 0 | D_i(x) > 0\} > 0$ *in the previous Proposition.*
When $r > 0$, $\beta \in (0, c_{i-1}(\theta_d^{i-1}))$ *and* $\alpha \in (0,1)$, *the function*

$$D_i(x) = \alpha(x-\beta)^+ + \left(\frac{1}{\sigma\beta}\right)^2 (r(K - \alpha\beta) + \eta) \left((x-\beta)^+\right)^2 e^{-\frac{x^2}{\eta}} \tag{3.11}$$

satisfies (A) *and the assumptions of Proposition 3.5.6 when* $\eta > 0$ *is small enough.*

Unfortunately, Proposition 3.5.6 does not apply to the simple dividend function $\alpha(x-\beta)^+$ *without addition of the second term in the right-hand-side of 3.11, even if from Figure 3.3(b) and the sentence following Lemma 3.5.2, one expects local monotonicity of the boundary.*

Proof. Let $0 \leq s < t$, $x > c_i(s)$ and τ be the smallest optimal stopping time for (s,x). We set $\bar{\tau} = \tau \mathbf{1}_{\{\tau < s\}} + \mathbf{1}_{\{\tau = s\}} \left(\inf\left\{v \geq s | \bar{S}_v^x \leq c_i(0)\right\} \wedge t\right)$. We have

$$u_i(t,x) - u_i(s,x) \geq \mathbb{E}\left[\mathbf{1}_{\{\tau = s\}}\left(e^{-r\bar{\tau}} u_i(t-\bar{\tau}, \bar{S}_{\bar{\tau}}^x) - e^{-rs} u_i(0, \bar{S}_s^x)\right)\right].$$

Since on $\{\tau = s\}$, $\bar{S}_s^x \geq c_i(0)$, on $\{\tau = s, \bar{\tau} < t\}$, $\bar{S}_{\bar{\tau}}^x = c_i(0)$, and $u_i(t-\bar{\tau}, c_i(0)) \geq (K - c_i(0)) = u_i(0, c_i(0))$. We then deduce that

$$u_i(t,x) - u_i(s,x) \geq \mathbb{E}\left[\mathbf{1}_{\{\bar{\tau} \geq s\}}\left(e^{-r\bar{\tau}} u_i(0, \bar{S}_{\bar{\tau}}^x) - e^{-rs} u_i(0, \bar{S}_s^x)\right)\right].$$

Applying Lemma 3.6.1, arguing like in the proof of Proposition 3.5.3 about the local martingale part and using that dv a.e. on $[s,t]$, $\bar{\tau} \geq v$ implies $\bar{S}_v^x > c_i(0)$, we get

$$u_i(t,x) - u_i(s,x) \geq \mathbb{E}\left[\int_s^t \mathbf{1}_{\{\bar{\tau} \geq v, \bar{S}_v^x > c_i(0)\}} e^{-rv} \left\{\begin{array}{l} -ru_i(0, \bar{S}_v^x) + r\bar{S}_v^x \partial_x u_{i-1}(\theta_d^{i-1}, \rho_i(\bar{S}_v^x))\rho_i'(\bar{S}_v^x) \\ +\frac{\sigma^2}{2}\left(\bar{S}_v^x \rho_i'\left(\bar{S}_v^x\right)\right)^2 \partial_{xx} u_{i-1}(\theta_d^{i-1}, \rho_i(\bar{S}_v^x)) \end{array}\right\} dv\right]$$
$$+ \frac{1}{2}\mathbb{E}\left[\int_s^t \int_{\mathbb{R}} \mathbf{1}_{\{\bar{\tau} \geq v\}} e^{-rv} \partial_x u_{i-1}(\theta_d^{i-1}, \rho_i(a)) \rho_i''(da) dL_v^a(\bar{S}^x)\right].$$

Like in the proof of Proposition 3.5.3, one checks that

$$\forall y \in (c_i(0), c_i(0)+\varepsilon],\ -ru_i(0,y) + ry\partial_x u_i(\theta_d^{i-1}, \rho_i(y))\rho_i'(y) + \frac{\sigma^2 y^2}{2} g_i(y) \geq \varepsilon$$
$$\forall y > c_i(0)+\varepsilon,\ -ru_i(0,y) + ry\partial_x u_i(\theta_d^{i-1}, \rho_i(y))\rho_i'(y) \geq -r(K+y),$$

$$\int_s^t \mathbf{1}_{\{\bar\tau \geq v\}} e^{-rv} \left(\bar S_v^x \rho_i'\left(\bar S_v^x\right)\right)^2 \partial_{xx} u_{i-1}(\theta_d^{i-1}, \rho_i(\bar S_v^x)) dv \geq -C_2 \int_s^t \mathbf{1}_{\left\{\bar\tau \geq v, \bar S_v^x > c_i(0)+\varepsilon\right\}} e^{-rv} \left(\bar S_v^x\right)^2 dv,$$

and that

$$\int_s^t \int_{\mathbb{R}} \mathbf{1}_{\{\bar\tau \geq v\}} e^{-rv} \partial_x u_{i-1}(\theta_d^{i-1}, \rho_i(a)) \rho_i''(da) dL_v^a(\bar S^x)$$
$$\geq \int_s^t \mathbf{1}_{\{\bar\tau \geq v\}} e^{-rv} \sigma^2 \left(\bar S_v^x\right)^2 \left[g_i(\bar S_v^x) \mathbf{1}_{\left\{\bar S_v^x \leq c_i(0)+\varepsilon\right\}} - C_1 (\bar S_v^x)^{C_1} \mathbf{1}_{\left\{\bar S_v^x > c_i(0)+\varepsilon\right\}}\right] dv.$$

Gathering all the inequalities, we get that it exists a finite constant $M \geq 0$ such that :

$$u_i(t,x) - u_i(s,x) \geq \int_s^t \left\{\mathbb{P}\left(\bar\tau \geq v\right) e^{-rv}\varepsilon - \mathbb{E}\left[\mathbf{1}_{\left\{\bar\tau \geq v,\ \bar S_v^x > c_i(0)+\varepsilon\right\}} M \left(1 + \left(\bar S_v^x\right)^{2+C_1}\right)\right]\right\} dv. \quad (3.12)$$

Applying Lemma 3.5.5 (with $p = 2 + C_1,\ t_1 = s$ and $\alpha = \frac{\varepsilon}{2}$), we obtain that for t small enough, uniformly for $x \leq c_i(0) + \frac{\varepsilon}{2}$, the right-hand-side of Equation 3.12 is non-negative.
With Proposition 3.4.1, we deduce the existence of $\eta > 0$ such that $\sup_{w \in [0,\eta]} c_i(w) \leq c_i(0) + \frac{\varepsilon}{2}$ and that

$$\forall 0 \leq s < t < \eta,\ \forall x \in (c_i(s), c_i(0) + \frac{\varepsilon}{2}),\ u_i(s,x) \leq u_i(t,x).$$

This inequality is still true for $x \leq c_i(s)$ since then $u_i(s,x) = (K-x)^+ \leq u_i(t,x)$.
Then, as soon as $0 \leq s < t < \eta$, $u_i(s, c_i(t)) \leq u_i(t, c_i(t)) = K - c_i(t)$ which implies that $c_i(t) \leq c_i(s)$.
□

Conclusion and further research

The continuity of the exercise boundary as well as the smooth contact property are likely to be generalized in a model with discrete dividends where the underlying asset price has a local volatility dynamics between the dividend dates with a positive local volatility function. We plan to investigate this extension in a future work. Assuming that the underlying stock price evolves as the exponential of some Lévy process between the dividend dates provides another natural generalization of the Black-Scholes model that could be considered (see [LM08] for the case without discrete dividends).

3.6 Appendix

3.6.1 Proof of Lemma 3.3.3

Proof. The existence of the right-hand limit at $c_i(\theta)$ for $\partial_x u_i(\theta,x)$ is an easy consequence of the second estimation. Since for $x < c_i(\theta)$, $\partial_{xx} u_i(\theta,x) = 0$ and for $x > c_i(\theta)$, by Proposition 3.2.4 and Lemma 3.2.2,

$$\begin{aligned}|\partial_{xx} u_i(\theta,x)| &= \left|\frac{2}{\sigma^2 x^2}\left(\partial_\theta u_i(\theta,x) + r u_i(\theta,x) - rx\partial_x u_i(\theta,x)\right)\right| \\ &\leq \frac{2}{\sigma^2 x^2}\left|\partial_\theta u_i(\theta,x)\right| + \frac{2r}{\sigma^2}\left(\frac{K}{x^2}+\frac{1}{x}\right),\end{aligned}$$

the second estimation is easily deduced from the first one. To prove the first estimation, we set

$$V_i : (\gamma,\nu,x) \mapsto \sup_{\tau\in[0,1]} \mathbb{E}\left[e^{-\gamma\frac{\nu^2}{2}\tau}\left(K - xe^{\frac{\nu^2}{2}(\gamma-1)\tau+\nu B_\tau}\right)^+ \mathbf{1}_{\{\tau<1\}} + e^{-\gamma\frac{\nu^2}{2}} u_i(0, xe^{\frac{\nu^2}{2}(\gamma-1)+\nu B_1})\mathbf{1}_{\{\tau=1\}}\right]$$

Because of the scaling property of the Brownian motion, for any positive $f : \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}$ and $\theta \in \mathbb{R}^+$, $\sup_{\tau\in[0,1]} \mathbb{E}\left[f(\theta\tau, \sqrt{\theta}B_\tau)\right] = \sup_{\tau\in[0,\theta]} \mathbb{E}\left[f(\tau, B_\tau)\right]$.
We deduce that $V_i\left(\frac{2r}{\sigma^2}, \sigma\sqrt{\theta}, x\right) = u_i(\theta,x)$ and

$$\limsup_{\theta'\to\theta}\left|\frac{u_i(\theta',x)-u_i(\theta,x)}{\theta'-\theta}\right| = \frac{\sigma}{2\sqrt{\theta}}\limsup_{\nu'\to\sigma\sqrt{\theta}}\left|\frac{V_i(\frac{2r}{\sigma^2},\nu',x) - V_i(\frac{2r}{\sigma^2},\sigma\sqrt{\theta},x)}{\nu'-\sigma\sqrt{\theta}}\right|.$$

Therefore it is enough to check that

$$\forall x,\nu\geq 0,\ \limsup_{\nu'\to\nu}\left|\frac{V_i(\gamma,\nu',x)-V_i(\gamma,\nu,x)}{\nu'-\nu}\right| \leq \nu\gamma\left(K+x\right) + x\left(\gamma\nu\left(2\mathcal{N}\left(\gamma\nu\right)-1\right) + 2\frac{e^{-\gamma^2\frac{\nu^2}{2}}}{\sqrt{2\pi}}\right). \quad (3.13)$$

Setting $(\gamma,\nu) = (\frac{2r}{\sigma^2}, \sigma\sqrt{\theta})$, the optimality of $\tau = \inf\left\{t\geq 0 | u_i(\theta-t, \bar{S}_t^x) + \bar{S}_t^x \leq K\right\} \wedge \theta$ for $u_i(\theta,x)$ translates into the optimality of

$$\tau^\star \stackrel{\text{def}}{=} \inf\left\{t\geq 0 | V_i(\gamma, \nu\sqrt{1-t}, xe^{\frac{\nu^2}{2}(\gamma-1)t+\nu B_t}) + xe^{\frac{\nu^2}{2}(\gamma-1)t+\nu B_t} \leq K\right\}\wedge 1$$

for $V_i(\gamma,\nu,x)$. This implies that

$$V_i(\gamma,\nu,x) + x = K\mathbb{E}\left[e^{-\frac{\nu^2}{2}\gamma\tau^\star}\right] + \mathbb{E}\left[\mathbf{1}_{\{\tau^\star=1\}}e^{-\frac{\nu^2}{2}\gamma}\left(u_i(0, xe^{\frac{\nu^2}{2}(\gamma-1)+\nu B_1}) + xe^{\frac{\nu^2}{2}(\gamma-1)+\nu B_1} - K\right)\right].$$

For any $\nu' \geq 0$, by definition of V_i,

$$V_i(\gamma,\nu',x) + x \geq K\mathbb{E}\left[e^{-\frac{\nu'^2}{2}\gamma\tau^\star}\right] + \mathbb{E}\left[\mathbf{1}_{\{\tau^\star=1\}}e^{-\frac{\nu'^2}{2}\gamma}\left(u_i(0, xe^{\frac{\nu'^2}{2}(\gamma-1)+\nu' B_1}) + xe^{\frac{\nu'^2}{2}(\gamma-1)+\nu' B_1} - K\right)\right].$$

Using that $x \mapsto x + u_i(0,x)$ is 1-lipschitz and non-decreasing by Lemma 3.2.2, then $u_i(0,.) \leq K$ and $(1-e^x)^+ \leq (-x)^+ \leq |x|$, one deduces

$$\begin{aligned}
V_i(\gamma,\nu',x)-V_i(\gamma,\nu,x) \geq & K\mathbb{E}\left[\left\{e^{-\frac{\nu'^2}{2}\gamma\tau^\star}-e^{-\frac{\nu^2}{2}\gamma\tau^\star}\right\}\mathbf{1}_{\{\tau^\star<1\}}\right] \\
& +\left[e^{-\frac{\nu'^2}{2}\gamma}-e^{-\frac{\nu^2}{2}\gamma}\right]\mathbb{E}\left[\mathbf{1}_{\{\tau^\star=1\}}\left(u_i(0,xe^{\frac{\nu^2}{2}(\gamma-1)+\nu B_1})+xe^{\frac{\nu^2}{2}(\gamma-1)+\nu B_1}\right)\right] \\
& -e^{-\frac{\nu'^2}{2}\gamma}\mathbb{E}\left[\mathbf{1}_{\{\tau^\star=1\}}xe^{\frac{\nu^2}{2}(\gamma-1)+\nu B_1}\left(1-e^{(\nu'-\nu)\left((\gamma-1)\frac{\nu+\nu'}{2}+B_1\right)}\right)^+\right]. \\
\geq & -K(e^{-\frac{\nu^2}{2}\gamma}-e^{-\frac{\nu'^2}{2}\gamma})^+(\mathbb{P}(\tau^\star<1)+\mathbb{P}(\tau^\star=1)) \\
& -x\left(1-e^{\frac{\nu^2-\nu'^2}{2}\gamma}\right)^+\mathbb{E}\left[\mathbf{1}_{\{\tau^\star=1\}}e^{-\frac{\nu^2}{2}+\nu B_1}\right] \\
& -e^{\frac{\nu^2-\nu'^2}{2}\gamma}\left|\nu'-\nu\right|\mathbb{E}\left[\mathbf{1}_{\{\tau^\star=1\}}xe^{-\frac{\nu^2}{2}+\nu B_1}\left|(\gamma-1)\frac{\nu+\nu'}{2}+B_1\right|\right] \\
\geq & -(K+x)\gamma\left|\nu-\nu'\right|\frac{\nu+\nu'}{2}-e^{\frac{\left|\nu^2-\nu'^2\right|}{2}\gamma}\left|\nu'-\nu\right|x\mathbb{E}\left[\left|(\gamma-1)\frac{\nu+\nu'}{2}+\nu+B_1\right|\right].
\end{aligned}$$

Remarking that for $y\in\mathbb{R}$, $\mathbb{E}|y+B_1|=y(2\mathcal{N}(y)-1)+\frac{2e^{-\frac{y^2}{2}}}{\sqrt{2\pi}}$ and combining the resulting inequality with the one deduced by exchanging ν and ν', we conclude that Equation 3.13 holds. □

3.6.2 Proofs of the auxiliary results of subsection 3.4.2

Proof of Lemma 3.4.8

Proof. Let $\theta>0$ and $x>c_i(\theta)$. For $a,b\in\mathbb{R}$ and $t\in[0,\theta]$, we define $Y_t^{a,b}=a+\frac{t}{\theta}(b-a)+\Xi_t$ where $(\Xi_s)_{s\in[0,\theta]}$ is a Brownian bridge on $[0,\theta]$ starting and ending at 0. Then $\left(Y_t^{a,b}\right)_{t\in[0,\theta]}$ is a Brownian bridge on $[0,\theta]$ starting at a and ending at b. For $y\geq 0$,

$$\begin{aligned}
\mathbb{P}\left(\tau=\theta|\bar{S}_\theta^x=y\right) &= \mathbb{P}\left(\forall t\in[0,\theta], Y_t^{0,\frac{1}{\sigma}\left(\ln\frac{y}{x}-\left(r-\frac{\sigma^2}{2}\right)\theta\right)}>\frac{1}{\sigma}\left(\ln\frac{c_i(\theta-t)}{x}-\left(r-\frac{\sigma^2}{2}\right)t\right)\right) \\
&= \mathbb{P}\left(\forall t\in[0,\theta],\quad \Xi_t>\frac{1}{\sigma}\left(\ln\frac{c_i(\theta-t)}{x}-\frac{t}{\theta}\ln\frac{y}{x}\right)\right)
\end{aligned}$$

and the monotonicity of $y\mapsto\mathbb{P}\left(\tau=\theta|\bar{S}_\theta^x=y\right)$ easily follows. For $y>K$, this implies

$$\frac{\mathbb{P}(\tau=\theta,\bar{S}_\theta^x\in(K,y))}{\mathbb{P}(\bar{S}_\theta^x\in(K,y))}\leq\mathbb{P}(\tau=\theta|\bar{S}_\theta^x=y).$$

Therefore, to prove the second assertion, we only need to check $\mathbb{P}(\tau=\theta,\bar{S}_\theta^x\in(K,y))>0$. Let $\eta=\inf\left\{t\geq 0|\bar{S}_t^x=\frac{y+K}{2}\right\}$. As $\sup_{t\geq 0}c_i(t)\leq K$, one has

$$\left\{\tau>\eta,\eta<\theta,\forall v\in[0,\theta-\eta]\bar{S}_{\eta+v}^x\in(K,y)\right\}\subset\left\{\tau=\theta,\bar{S}_\theta^x\in(K,y)\right\}.$$

By the strong Markov property and the continuity of the Black-Scholes model, one deduces

$$\begin{aligned}\mathbb{P}(\tau=\theta,\bar{S}^x_\theta\in(K,y)) &\geq \mathbb{E}\left[\mathbf{1}_{\{\tau>\eta,\eta<\theta\}}\mathbb{P}(\forall v\in[0,t],\bar{S}_v^{\frac{y+K}{2}}\in(K,y))\Big|_{t=\theta-\eta}\right]\\ &\geq \mathbb{P}(\tau>\eta,\eta<\theta)\mathbb{P}(\forall v\in[0,\theta],\bar{S}_v^{\frac{y+K}{2}}\in(K,y))\\ &\geq \mathbb{P}\left(\tau=\theta,\bar{S}^x_\theta\geq y\right)\mathbb{P}(\forall v\in[0,\theta],\bar{S}_v^{\frac{y+K}{2}}\in(K,y)).\end{aligned}$$

The last factor in the right-hand-side is positive. By comonotony,

$$\mathbb{P}\left(\tau=\theta,\bar{S}^x_\theta\geq y\right)=\mathbb{E}\left[\mathbb{P}(\tau=\theta|\bar{S}^x_\theta)\mathbf{1}_{\{\bar{S}^x_\theta\geq y\}}\right]\geq\mathbb{P}\left(\tau=\theta\right)\mathbb{P}\left(\bar{S}^x_\theta\geq y\right).$$

One concludes by remarking that

$$K\mathbb{E}\left[e^{-r\tau}\right]-x+\mathbb{E}\left[e^{-r\theta}\mathbf{1}_{\{\tau=\theta\}}\left(u_i(0,\bar{S}^x_\theta)+\bar{S}^x_\theta-K\right)\right]=u_i(\theta,x)>K-x\geq K\mathbb{E}\left[e^{-r\tau}\right]-x$$

implies positivity of $\mathbb{P}(\tau=\theta)$. □

Proof of Lemma 3.4.9

Proof. Let $\theta,\epsilon>0$, $x\geq 0$ and τ be an optimal stopping time for $u_i(\theta,x)$. Since

$$u_i(\theta,x+\epsilon)\geq\mathbb{E}\left[e^{-r\tau}\left(K-\bar{S}^{x+\epsilon}_\tau\right)^+\mathbf{1}_{\{\tau<\theta\}}+e^{-r\theta}u_i(0,\bar{S}^{x+\epsilon}_\theta)\mathbf{1}_{\{\tau=\theta\}}\right]$$

and $(K-\bar{S}^{x+\epsilon}_\tau)^+-(K-\bar{S}^x_\tau)^+\geq\bar{S}^x_\tau-\bar{S}^{x+\epsilon}_\tau$, we have

$$\begin{aligned}\tfrac{u_i(\theta,x+\epsilon)-u_i(\theta,x)}{\epsilon} &\geq \tfrac{1}{\epsilon}\mathbb{E}\left[e^{-r\tau}\left(\bar{S}^x_\tau-\bar{S}^{x+\epsilon}_\tau\right)\mathbf{1}_{\{\tau<\theta\}}+e^{-r\theta}\left(u_i(0,\bar{S}^{x+\epsilon}_\theta)-u_i(0,\bar{S}^x_\theta)\right)\mathbf{1}_{\{\tau=\theta\}}\right]\\ &= -\mathbb{E}\left[e^{-r\tau}\bar{S}^1_\tau\mathbf{1}_{\{\tau<\theta\}}\right]+\ \mathbb{E}\left[e^{-r\theta}\bar{S}^1_\theta\frac{u_i(0,\bar{S}^{x+\epsilon}_\theta)-u_i(0,\bar{S}^x_\theta)}{\bar{S}^{x+\epsilon}_\theta-\bar{S}^x_\theta}\mathbf{1}_{\{\tau=\theta\}}\right]\\ &= -\mathbb{Q}\left(\tau<\theta\right)+\ \mathbb{E}^{\mathbb{Q}}\left[\frac{u_i(0,\bar{S}^{x+\epsilon}_\theta)-u_i(0,\bar{S}^x_\theta)}{\bar{S}^{x+\epsilon}_\theta-\bar{S}^x_\theta}\mathbf{1}_{\{\tau=\theta\}}\right]\\ &= -1+\ \mathbb{E}^{\mathbb{Q}}\left[\left(1+\frac{u_i(0,\bar{S}^{x+\epsilon}_\theta)-u_i(0,\bar{S}^x_\theta)}{\bar{S}^{x+\epsilon}_\theta-\bar{S}^x_\theta}\right)\mathbf{1}_{\{\tau=\theta\}}\right]\end{aligned}$$

where we used $\bar{S}^x_\theta=x\bar{S}^1_\theta$ for the first equality and $\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}\Big|_{\mathcal{F}_\theta}=e^{-r\theta}\bar{S}^1_\theta$ for the second one. The first assertion is deduced by dominated convergence using that, according to Lemma 3.2.2, $x\mapsto u_i(0,x)$ is 1-Lipschitz and therefore almost surely differentiable,.

The smallest optimal stopping time for $u_i(\theta,x+\epsilon)$ is $\tau^\epsilon=\theta\wedge\inf\left\{t\in[0,\theta]|\bar{S}^{x+\epsilon}_t\leq c_i(\theta-t)\right\}$. Clearly, $\mathbb{P}$-almost surely, for any $\epsilon>\epsilon'$, $\tau^\epsilon\geq\tau^{\epsilon'}$ and one may define $\overline{\tau}$ as $\lim_{\epsilon\to 0_+}\tau^\epsilon$. Moreover, $\overline{\tau}\geq\tau^\star$ where $\tau^\star$ is the smallest optimal stopping time for $u_i(\theta,x)$. As $(\mathcal{F}_t)_t$ is a right-continuous filtration, $\overline{\tau}$ is a stopping time (cf (4.17) p.46 of [RY91]). By optimality of τ^ϵ,

$$u_i(\theta, x+\epsilon) = \mathbb{E}\left[e^{-r\tau^\epsilon}\right] K - (x+\epsilon) + \mathbb{E}\left[e^{-r\theta}\mathbf{1}_{\{\tau^\epsilon=\theta\}}\left(u_i\left(0, \bar{S}_\theta^{x+\epsilon}\right) + \bar{S}_\theta^{x+\epsilon} - K\right)\right].$$

Since $x \mapsto x + u_i(0,x)$ is 1-Lipschitz, one may take the limit $\epsilon \to 0$ in this equality and obtain

$$u_i(\theta, x) = \mathbb{E}\left[e^{-r\overline{\tau}}\right] K - x + \mathbb{E}\left[e^{-r\theta}\mathbf{1}_{\{\overline{\tau}=\theta\}}\left(u_i\left(0, \bar{S}_\theta^{x}\right) + \bar{S}_\theta^{x} - K\right)\right],$$

which implies that $\overline{\tau}$ is also an optimal stopping time for $u_i(\theta, x)$.
When $\tau^\epsilon < \theta$, $\bar{S}_{\tau^\epsilon}^x \leq \bar{S}_{\tau^\epsilon}^{x+\epsilon} \leq K$. Therefore

$$\begin{aligned}\tfrac{u_i(\theta,x+\epsilon)-u_i(\theta,x)}{\epsilon} &\leq \tfrac{1}{\epsilon}\mathbb{E}\left[e^{-r\tau^\epsilon}\left(\bar{S}_{\tau^\epsilon}^x - \bar{S}_{\tau^\epsilon}^{x+\epsilon}\right)\mathbf{1}_{\{\tau^\epsilon<\theta\}} + e^{-r\theta}\left(u_i(0,\bar{S}_\theta^{x+\epsilon}) - u_i(0,\bar{S}_\theta^x)\right)\mathbf{1}_{\{\tau^\epsilon=\theta\}}\right] \\ &= -1 + \ \mathbb{E}^{\mathbb{Q}}\left[\left(1 + \frac{u_i(0,\bar{S}_\theta^{x+\epsilon}) - u_i(0,\bar{S}_\theta^x)}{\bar{S}_\theta^{x+\epsilon} - \bar{S}_\theta^x}\right)\mathbf{1}_{\{\tau^\epsilon=\theta\}}\right].\end{aligned}$$

Letting $\epsilon \to 0$ in this inequality, we obtain by dominated convergence

$$\partial_x u_i(\theta, x) + 1 \leq \mathbb{E}^{\mathbb{Q}}\left[\mathbf{1}_{\{\overline{\tau}=\theta\}}\left(1 + \partial_x u_i(0, \bar{S}_\theta^x)\right)\right],$$

which concludes the proof. □

3.6.3 Proofs of the auxiliary results of Section 3.5

Proof of Lemma 3.5.2

Proof. Let $\theta > 0$. Using the definition of u_i, Equation 3.4, and the Cauchy Schwarz inequality, we get

$$\begin{aligned}u_i(\theta,x) \geq& Ke^{-r\theta} - x + e^{-r\theta}\mathbb{E}\left[u_i(0,\bar{S}_\theta^x) + \bar{S}_\theta^x - K\right] \\ \geq& Ke^{-r\theta} - x + e^{-r\theta}\mathbb{E}\left[\alpha\left|(\bar{S}_\theta^x - c_i(0))^+\right|^\beta\right] \\ &\quad + e^{-r\theta}\mathbb{E}\left[\mathbf{1}_{\{\bar{S}_\theta^x \notin V\}}\left(u_i(0,\bar{S}_\theta^x) + \bar{S}_\theta^x - K - \alpha\left|\left(\bar{S}_\theta^x - c_i(0)\right)^+\right|^\beta\right)\right] \\ \geq& Ke^{-r\theta} - x + e^{-r\theta}\mathbb{E}\left[\alpha\left|(\bar{S}_\theta^x - c_i(0))^+\right|^\beta\right] - e^{-r\theta}\mathbb{E}\left[\mathbf{1}_{\{\bar{S}_\theta^x \notin V\}}\alpha\left|\bar{S}_\theta^x\right|^\beta\right] \\ \geq& Ke^{-r\theta} - x + e^{-r\theta}\mathbb{E}\left[\alpha\left|(\bar{S}_\theta^x - c_i(0))^+\right|^\beta\right] - \alpha e^{-r\theta}x^\beta e^{\beta\left(r-\frac{\sigma^2}{2}\right)\theta+\beta^2\sigma^2\theta}\sqrt{\mathbb{P}(\bar{S}_\theta^x \notin V)}.\end{aligned}$$

Let $\epsilon > 0$ be such that $(c_i(0) - 2\epsilon, c_i(0) + 2\epsilon) \subset V$. For $x \in (c_i(0) - \epsilon, c_i(0) + \epsilon)$,

$$\mathbb{P}\left(\bar{S}_\theta^x \notin V\right) \leq \mathbb{P}\left(\bar{S}_\theta^x \notin (x-\epsilon, x+\epsilon)\right) \leq 2\mathcal{N}\left\{\frac{1}{\sigma\sqrt{\theta}}\left(\left(r + \frac{\sigma^2}{2}\right)\theta + \ln\max\left(\frac{x-\epsilon}{x}, \frac{x}{x+\epsilon}\right)\right)\right\}.$$

We deduce that

$$u_i(\theta,x) \geq Ke^{-r\theta} - x + e^{-r\theta}\mathbb{E}\left[\alpha\left|(\bar{S}^x_\theta - c_i(0))^+\right|^\beta\right] + o(\theta), \tag{3.14}$$

where the term $o(\theta)$ is uniform for $x \in (c_i(0)-\epsilon, c_i(0)+\epsilon)$. In order to bound the third term of the right-hand-side from below, we first deal with $\phi(\theta) \stackrel{\text{def}}{=} \mathbb{E}\left[\left|\left(\bar{S}^1_\theta - 1\right)^+\right|^\beta\right]$. Using the change of variables $z = \sigma\sqrt{\theta}u$ for the second equality, we have

$$\begin{aligned}
\phi(\theta) &= \int_0^{+\infty} z^\beta e^{-\frac{1}{2\sigma^2\theta}\left(\ln(1+z)-\left[r-\frac{\sigma^2}{2}\right]\theta\right)^2} \frac{dz}{\sqrt{2\pi\theta}\sigma(1+z)} \\
&\geq e^{-\left[\frac{r}{\sigma}-\frac{\sigma}{2}\right]^2\theta} \int_0^{+\infty} z^\beta e^{-\frac{1}{\sigma^2\theta}\ln^2(1+z)} \frac{dz}{\sqrt{2\pi\theta}\sigma(1+z)} \\
&\geq e^{-\left[\frac{r}{\sigma}-\frac{\sigma}{2}\right]^2\theta} \int_0^{+\infty} z^\beta e^{-\frac{z^2}{\sigma^2\theta}} \frac{dz}{\sqrt{2\pi\theta}\sigma(1+z)} = e^{-\left[\frac{r}{\sigma}-\frac{\sigma}{2}\right]^2\theta}\sigma^\beta\theta^{\frac{\beta}{2}} \int_0^{+\infty} \frac{u^\beta e^{-u^2}du}{\sqrt{2\pi}(1+u\sigma\sqrt{\theta})} \\
&\geq e^{-\left[\frac{r}{\sigma}-\frac{\sigma}{2}\right]^2\theta}\sigma^\beta\theta^{\frac{\beta}{2}} \int_0^{+\infty} \frac{u^\beta e^{-u^2}}{\sqrt{2\pi}}\left(1-u\sigma\sqrt{\theta}\right)du \\
&= e^{-\left[\frac{r}{\sigma}-\frac{\sigma}{2}\right]^2\theta}\sigma^\beta\theta^{\frac{\beta}{2}} \frac{1}{\sqrt{8\pi}}\left[\Gamma\left(\frac{1+\beta}{2}\right) - \sigma\sqrt{\theta}\Gamma\left(\frac{3+\beta}{2}\right)\right] \\
&= e^{-\left[\frac{r}{\sigma}-\frac{\sigma}{2}\right]^2\theta}\sigma^\beta\theta^{\frac{\beta}{2}} \frac{1}{\sqrt{8\pi}}\Gamma\left(\frac{1+\beta}{2}\right)\left[1-\sigma\sqrt{\theta}\frac{1+\beta}{2}\right].
\end{aligned}$$

Thus, for $\theta < \frac{1}{\sigma^2(1+\beta)^2}$ and $C = \frac{1}{2}e^{-\frac{\left(\frac{r}{\sigma}-\frac{\sigma}{2}\right)^2}{\sigma^2(1+\beta)^2}}\frac{\sigma^\beta}{\sqrt{8\pi}}\Gamma\left(\frac{1+\beta}{2}\right)$, one has $\phi(\theta) \geq C\theta^{\frac{\beta}{2}}$.
Let $x < c_i(0)$ and $\tau = \inf\left\{t \geq 0 | \bar{S}^x_t \geq c_i(0)\right\}$. For $\theta < \frac{1}{\sigma^2(1+\beta)^2}$, using the strong Markov property then Formula 2.0.2 p.223 [BS96], one has

$$\begin{aligned}
\mathbb{E}\left[\left|\left(\bar{S}^x_\theta - c_i(0)\right)^+\right|^\beta\right] &= |c_i(0)|^\beta\,\mathbb{E}\left[\mathbb{E}\left[\left|\left(\bar{S}^1_{\theta-\tau}-1\right)^+\right|^\beta |\mathcal{F}_\tau\right]\mathbf{1}_{\{\tau<\theta\}}\right] \\
&= |c_i(0)|^\beta\,\mathbb{E}\left[\phi\left(\theta-\tau\right)\mathbf{1}_{\{\tau<\theta\}}\right] \\
&\geq |c_i(0)|^\beta\, C\theta^{\frac{\beta}{2}}\mathbb{E}\left[\left(1-\frac{\tau}{\theta}\right)^{\frac{\beta}{2}}\mathbf{1}_{\{\tau<\theta\}}\right] \\
&\geq |c_i(0)|^\beta\, C\theta^{\frac{\beta}{2}}\frac{1}{\sigma}\ln\frac{c_i(0)}{x}\int_0^\theta\left(1-\frac{t}{\theta}\right)^{\frac{\beta}{2}}\frac{1}{\sqrt{2\pi t^3}}e^{-\frac{1}{2\sigma^2 t}\left(\left(\frac{\sigma^2}{2}-r\right)t+\ln\frac{c_i(0)}{x}\right)^2}dt \\
&\geq |c_i(0)|^\beta\; e^{\frac{1}{2\sigma^2}\left[2\left(\frac{\sigma^2}{2}-r\right)\ln\frac{x}{c_i(0)}-\left(\frac{\sigma^2}{2}-r\right)^2\theta\right]}\,C\theta^{\frac{\beta}{2}} \\
&\qquad\times\underbrace{\frac{1}{\sigma\sqrt{2\pi\theta}}\ln\frac{c_i(0)}{x}\int_0^1(1-u)^{\frac{\beta}{2}}\frac{1}{\sqrt{u^3}}e^{-\frac{1}{2\sigma^2\theta u}\ln^2\frac{c_i(0)}{x}}du}_{:=\psi(\theta,x)}.
\end{aligned}$$

Hence

$$\exists M,\eta>0,\ \forall\,(\theta,x)\in(0,\eta)\times(c_i(0)e^{-\sigma\theta^{\frac{1}{3}}},c_i(0)),\ \mathbb{E}\left[\left|\left(\bar{S}^x_\theta-c_i(0)\right)^+\right|^\beta\right] \geq M\theta^{\frac{\beta}{2}}\psi(\theta,x). \tag{3.15}$$

Setting $\gamma(x) = \frac{1}{\sigma\sqrt{\theta}} \ln \frac{c_i(0)}{x}$, we have $\psi(\theta,x) = \frac{\gamma(x)}{\sqrt{2\pi}} \int_0^1 (1-u)^{\frac{\beta}{2}} u^{-\frac{3}{2}} e^{-\frac{\gamma^2(x)}{2u}} du$. With the change of variables $t = \frac{1}{u} - 1$, we deduce that $\psi(\theta,x) = \frac{\gamma(x)}{\sqrt{2\pi}} e^{-\frac{\gamma^2(x)}{2}} \Gamma\left(\frac{\beta}{2}+1\right) \mathrm{U}\left(\frac{\beta}{2}+1; \frac{3}{2}; \frac{\gamma^2(x)}{2}\right)$ where $\mathrm{U}(a,b,z) = \frac{1}{\Gamma(a)} \int_0^{+\infty} e^{-tz} t^{a-1}(1+t)^{b-a-1} dt$ is the confluent hypergeometric function of the second kind. By 13.5.2 p.504 [AS72],

$$\text{for } z \to +\infty,\ \mathrm{U}\left(\frac{\beta}{2}+1; \frac{3}{2}; z\right) = z^{-(\frac{\beta}{2}+1)}(1+O(1/z)).$$

Then we choose θ small enough to ensure that $x(\theta) = c_i(0)e^{-\sigma\sqrt{\theta((2-\beta)|\ln\theta| - (\delta+\beta)\ln|\ln\theta|)}}$ is well defined. Since $\gamma(x(\theta)) = \sqrt{(2-\beta)|\ln\theta| - (\delta+\beta)\ln|\ln\theta|}$ tends to $+\infty$ as $\theta \to 0$, we deduce

$$\begin{aligned}\psi(\theta,x(\theta)) =& \frac{\Gamma\left(\frac{\beta}{2}+1\right) 2^{1+\frac{\beta}{2}}}{((2-\beta)|\ln\theta| - (\delta+\beta)\ln|\ln\theta|)^{\frac{\beta+1}{2}} \sqrt{2\pi}} \theta^{1-\frac{\beta}{2}} |\ln\theta|^{\frac{\delta+\beta}{2}} \left(1 + O\left(\frac{1}{|\ln\theta|}\right)\right) \\ =& \frac{\Gamma\left(\frac{\beta}{2}+1\right) 2^{1+\frac{\beta}{2}}}{\sqrt{2\pi}(2-\beta)^{\frac{\beta+1}{2}}} \theta^{1-\frac{\beta}{2}} |\ln\theta|^{\frac{\delta-1}{2}} \left(1 + O\left(\frac{\ln|\ln\theta|}{|\ln\theta|}\right)\right).\end{aligned}$$

Plugging this into Equation 3.15, we conclude that it exists a constant $\kappa > 0$ such that as $\theta \to 0$,

$$\mathbb{E}\left[\left|\left(\bar{S}_\theta^{x(\theta)} - c_i(0)\right)^+\right|^\beta\right] \geq \kappa\theta|\ln\theta|^{\frac{\delta-1}{2}} \left(1 + O\left(\frac{\ln|\ln\theta|}{|\ln\theta|}\right)\right).$$

With Equation 3.14, this implies that

$$u_i(\theta,x(\theta)) \geq K - x(\theta) + \theta\left(\kappa|\ln\theta|^{\frac{\delta-1}{2}} - rK\right) + o(\theta)$$

and the conclusion follows by positivity of the factor $\kappa|\ln\theta|^{\frac{\delta-1}{2}} - rK$ for θ small enough. □

Proof of Lemma 3.5.5.

Proof. Ideas are similar to those of the proof of Proposition 2.2 of [JV11]. For $\alpha > 0$, according to Proposition 3.4.1, there exists $\eta > 0$ such that $\sup_{w\in[0,\eta]} c_i(w) \leq c_i(0) + \frac{\alpha}{2}$. Let us suppose that $t_1 \in [0,\eta]$. Let $x \leq c_i(0)+\alpha$ and $v \geq 0$.
Setting $\tilde{\tau} = \inf\left\{w \geq 0 | \bar{S}_w^x \geq c_i(0)+\alpha\right\}$, we have

$$\mathbf{1}_{\{\tau\geq v\}} \geq \mathbf{1}_{\left\{\tau\geq\tilde{\tau},\ \tilde{\tau}\leq v,\ \forall w\in[\tilde{\tau},v],\ \bar{S}_w^x > c_i(0)+\alpha\right\}}$$

Using the strong Markov property, we deduce that

$$\mathbb{P}(\tau \geq v) \geq \mathbb{P}(\tau\geq\tilde{\tau},\ \tilde{\tau}\leq v)\,\mathbb{P}\left(\inf_{w\in[0,v]} \bar{S}_w^1 > \frac{c_i(0)+\frac{\alpha}{2}}{\hat{c}_i(0)+\alpha}\right). \tag{3.16}$$

Whereas, by continuity of the trajectories of $\bar{S}^x$ and since $x \leq c_i(0)+\alpha$,

$$\mathbf{1}_{\left\{\tau\geq v,\ \bar{S}^x_v\geq c_i(0)+2\alpha\right\}}\leq\mathbf{1}_{\left\{\tau\geq\tilde{\tau},\tilde{\tau}\leq v,\ \bar{S}^x_v\geq c_i(0)+2\alpha\right\}}.$$

Again by the strong Markov property, we deduce that

$$\mathbb{E}\left[\left(\bar{S}^x_v\right)^p\mathbf{1}_{\left\{\tau\geq v,\ \bar{S}^x_v\geq c_i(0)+2\alpha\right\}}\right]\leq\mathbb{E}\left[\mathbf{1}_{\{\tau\geq\tilde{\tau},\tilde{\tau}\leq v\}}\left(c_i(0)+\alpha\right)^p\mathbb{E}\left[\left(\bar{S}^1_w\right)^p\mathbf{1}_{\left\{\bar{S}^1_w\geq\frac{c_i(0)+2\alpha}{c_i(0)+\alpha}\right\}}\right]_{w=v-\tilde{\tau}}\right].\tag{3.17}$$

Then by defining $\tilde{\mathbb{P}}$ as $\left.\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}}\right|_{\mathcal{F}_t}=e^{p\sigma B_t-\frac{p^2\sigma^2 t}{2}}$, we get

$$\mathbb{E}\left[\left(\bar{S}^x_v\right)^p\mathbf{1}_{\left\{\tau\geq v,\ \bar{S}^x_v\geq c_i(0)+2\alpha\right\}}\right]\leq\mathbb{P}\left(\tau\geq\tilde{\tau},\ \tilde{\tau}\leq v\right)\left(c_i(0)+\alpha\right)^p e^{\left(pr+\sigma^2\frac{p(p-1)}{2}\right)v}\sup_{0\leq w\leq v}\tilde{\mathbb{P}}\left(\bar{S}^1_w\geq\frac{c_i(0)+2\alpha}{c_i(0)+\alpha}\right).\tag{3.18}$$

Notice that for any $t,x,y\geq 0$, $\mathbb{P}(\bar{S}^x_t\geq y)\leq\tilde{\mathbb{P}}(\bar{S}^x_t\geq y)$. So, we deduce that

$$\frac{\mathbb{E}\left[\left(1+\left(\bar{S}^x_v\right)^p\right)\mathbf{1}_{\left\{\tau\geq v,\ \bar{S}^x_v\geq c_i(0)+2\alpha\right\}}\right]}{\mathbb{P}(\tau\geq v)}\leq\frac{\left(1+\left(c_i(0)+\alpha\right)^p e^{\left(pr+\sigma^2\frac{p(p-1)}{2}\right)v}\right)\sup_{0\leq w\leq v}\tilde{\mathbb{P}}\left(\bar{S}^1_w\geq\frac{c_i(0)+2\alpha}{c_i(0)+\alpha}\right)}{\mathbb{P}\left(\inf_{w\in[0,v]}\bar{S}^1_w>\frac{c_i(0)+\frac{\alpha}{2}}{\hat{c}_i(0)+\alpha}\right)}.\tag{3.19}$$

This concludes the proof since when v tends to 0, the numerator tends to 0 whereas the denominator tends to 1. □

Itô Tanaka formula

Lemma 3.6.1 *For $i\geq 1$, assume that D_i is difference of two convex functions. Then*

$$\begin{aligned}du_{i-1}(\theta_d^{i-1},\rho_i(\bar{S}^x_t))=&\partial_x u_{i-1}(\theta_d^{i-1},\rho_i(\bar{S}^x_t))\rho_i'(\bar{S}^x_t)d\bar{S}^x_t+\frac{1}{2}\int_{\mathbb{R}}\partial_x u_{i-1}(\theta_d^{i-1},\rho_i(a))dL^a_t(\bar{S}^x)\rho_i''(da)\\&+\frac{1}{2}\partial_{xx}u_{i-1}(\theta_d^{i-1},\rho_i(\bar{S}^x_t))\left(\rho_i'(\bar{S}^x_t)\right)^2 d\left\langle\bar{S}^x\right\rangle_t\end{aligned}$$

Proof. By the Itô-Tanaka formula,

$$d\rho_i(\bar{S}^x_t)=\rho_i'(\bar{S}^x_t)d\bar{S}^x_t+\frac{1}{2}\int_{\mathbb{R}}dL^a_t(\bar{S}^x)\rho_i''(da).$$

Hence $X_t=\rho_i(\bar{S}^x_t)$ is a continuous semi-martingale with bracket $\langle X\rangle_t=\int_0^t\left(\rho_i'(\bar{S}^x_s)\right)^2 d\left\langle\bar{S}^x\right\rangle_s$. By Lemma 3.3.3, since $\theta_d^{i-1}>0$, the function $f(x)=\partial_{xx}u_{i-1}\left(\theta_d^{i-1},\bullet\right)$ is bounded. The next Lemma ensures that

$$\begin{aligned}du_{i-1}(\theta_d^{i-1},\rho_i(\bar{S}^x_t))=&\partial_x u_{i-1}(\theta_d^{i-1},\rho_i(\bar{S}^x_t))\left(\rho_i'(\bar{S}^x_t)d\bar{S}^x_t+\frac{1}{2}\int_{\mathbb{R}}\rho_i''(da)dL^a_t(\bar{S}^x)\right)\\&+\frac{1}{2}\partial_{xx}u_{i-1}(\theta_d^{i-1},\rho_i(\bar{S}^x_t))\left(\rho_i'(\bar{S}^x_t)\right)^2 d\left\langle\bar{S}^x\right\rangle_t.\end{aligned}$$

One concludes since, by Proposition 1.3 p.222 [RY91], $\mathbb{P} \otimes |\rho_i''|(da)$ a.e., the measure $dL_t^a(\bar{S}^x)$ is supported by $\{t : \bar{S}_t^x = a\}$.

□

Lemma 3.6.2 *Let X be a continuous semi-martingale and f a $\mathcal{C}^1$ function, $\mathcal{C}^2$ on $[0, x^\star)$ and $(x^\star, +\infty)$, such that either $\inf_{x\in\mathbb{R}} f''(x)$ or $\sup_{x\in\mathbb{R}} f''(x)$ is finite. Then, almost surely,*

$$\int_0^t \mathbf{1}_{\{X_s = x^\star\}} d\left\langle X \right\rangle_s = 0 \text{ and } f(X_t) = f(X_0) + \int_0^t f'(X_s) dX_s + \frac{1}{2} \int_0^t f''(X_s) d\left\langle X \right\rangle_s .$$

Proof. The first assertion is a consequence of the occupation times formula and ensures that differentiability of f' at $x^\star$ is not needed for the right-hand-side of the second equality to be well defined. By hypothesis, it exists $0 \leq M < \infty$ such that either $x \mapsto f(x) + Mx^2$ or $x \mapsto f(x) - Mx^2$ is convex and consequently f is the difference of two convex functions. So we can apply the Itô-Tanaka formula and conclude by the occupation times formula. □

4

Regularity of the American put option in the Lévy's exponential model with general discrete dividends

Summary. We analyze how we can approximate the value function near the dividend dates. We analyze the regularity of the value function and of the optimal exercise boundary of the American Put option when the underlying asset pays a discrete dividend at known times during the lifetime of the option. The ex-dividend asset price process is assumed to follow the exponential Lévy dynamics and the dividend amount is a deterministic function of the ex-dividend asset price just before the dividend date. This function is assumed to be non-negative, non-decreasing and with growth rate not greater than 1. Under some explicit conditions, we prove that the exercise boundary is continuous at any time but the dividend dates.

Introduction

This chapter is organized as follows. In the first section, we give the notations and assumptions used in the following sections. The second section is devoted to the statement of the American option problem in the framework of discrete dividends. We will assume that its value is given by the Snell envelop of some process, and we will state some first results derived from [JJ12] and [LM08], especially the fact that the exercise region is fully characterized by a time-dependent curve called the exercise boundary. In the third section, we give conditions for this last one to be continuous. A necessary and sufficient condition for the smooth fit to hold is given in the fourth section. The last section deals with a numerical procedure to get the price of the American option in our model.

4.1 Notations and assumptions

4.1.1 Notations

N_1. $\Re(z)$ denotes the real part of the complex number $z \in \mathbb{C}$.

N_2. $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space.

N$_3$. X is a real Lévy process with characteristic exponent defined on some band of the complex plane denoted by $\Lambda = \{\lambda \in \mathbb{C} : \Re(\lambda) \in \mathrm{I}\}$ where I is some interval of the real line containing $[0,1]$:

$$\psi : \lambda \in \Lambda \mapsto \psi(\lambda) = \frac{\sigma^2}{2}\lambda^2 + \gamma\lambda + \int_{\mathbb{R}} \left(e^{\lambda y} - 1 - \lambda y \mathbf{1}_{\{|y|\le 1\}}\right) \nu(dy) \tag{4.1}$$

In particular, $\mathbb{E}\left[e^{\lambda X_t}\right] = e^{t\psi(\lambda)}$. We assume that $\psi(1) = r \ge 0$ and it enables us to define the probability measure $\mathbb{Q}$ (of change of numeraire) by

$$\left.\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}\right|_{\sigma\left(\left. X_s\right|_{s\le t}\right)} = e^{X_t - rt}.$$

N$_4$. $\bar{S}_t^x \overset{\text{def}}{=} xe^{X_t}$ is the standard spot price process in a standard Lévy's exponential model, note that $e^{-rt}\bar{S}_t^x$ is a martingale under $\mathbb{P}$.

N$_5$. We denote by $\mathcal{L}$ the infinitesimal generator of X which satisfies for a function f smooth enough :

$$\mathcal{L}f(x) = \frac{\sigma^2}{2} f''(x) + \gamma f'(x) + \int_{\mathbb{R}} \left(f(x+y) - f(x) - y\mathbf{1}_{\{|y|\le 1\}} f'(x)\right) \nu(dy) \tag{4.2}$$

N$_6$. For a process Y, $\mathcal{F}^Y_{\le\theta}$ denotes the set of all the stopping times of the filtration $\mathcal{F}^Y$ generated by $(Y_t)_{t\ge 0}$ which are not greater than θ.

4.1.2 Assumptions

A$_1$. A function D is said to be a *natural dividend function* when :

$$\begin{cases} \text{(a) } D \text{ is non-decreasing and non-negative,} \\ \text{(b) } \rho : x \mapsto x - D(x) \text{ is non-decreasing and non-negative.} \end{cases}$$

4.2 American option

4.2.1 Statement of the problem

We consider the American put option with maturity $T > 0$ and strike $K > 0$ written on an underlying stock S. We assume that the stochastic dynamics of the ex-dividend price process of this stock can be modelled by the exponential Lévy model which generalizes [JJ12] where the dynamics is given by the Black-Scholes model. However, we consider the same modelling for the payment of dividends. At deterministic times $0 \le t_d^I < t_d^{I-1} < \cdots < t_d^i < \cdots < t_d^1 < T$, this stock is paying discrete dividends. At each dividend time t_d^i, the value of the stock becomes $S_{t_d^i} = S_{t_d^i-} - D_i\left(S_{t_d^i-}\right)$ where $D_i(S_{t_d^i-})$ is the value of the dividend payment. For each $i \in \{1, \ldots, I\}$, the function D_i is supposed to be a *natural dividend function*, and we define $\rho_i(x) = x - D_i(x)$. We then abusively refer to A_1 as the assumption that each dividend function is natural and we will assume that this assumption is always fulfilled in the next sections.

In the financial market which consists of both financial assets S and cash discounted at rate r, we know at least in the Black-Scholes model (cf [Myn92]) that

$$P_t = \operatorname{ess} \sup_{\tau \in \mathcal{F}^S : t \leq \tau \leq T} \mathbb{E}\left[e^{-r(T-t)}(K - S_T)^+ | \mathcal{F}_t^S\right]$$

is an arbitrage-free value. **From now on** we state that the price of the American put option is P_t. Following the construction of [JJ12], we build recursively the functions u_i. We recall that u_0 is the pricing function of the standard American put option with strike K, in the Lévy's exponential model described in Section 4.1. For results on u_0, we will refer to [LM08]. We then set $t_d^0 = T$, and for $i \geq 0$, $\theta_d^i = t_d^i - t_d^{i+1}$, in order to define for $i \geq 1$:

$$u_i : (\theta, x) \in \mathbb{R}_+^2 \mapsto \sup_{\tau \in \mathcal{F}_{\leq \theta}^X} \mathbb{E}\left[e^{-r\tau}\left(\left(K - xe^{X_\tau}\right)^+ \mathbf{1}_{\{\tau < \theta\}} + \mathbf{1}_{\{\tau = \theta\}} u_{i-1}(\theta_d^{i-1}, \rho_i(xe^{X_\theta}))\right)\right] \tag{4.3}$$

4.2.2 First results

We insist on the fact that notations are the same as in [JJ12] up to the fact that $\bar{S}_t^x = xe^{X_t}$.We state the results of [JJ12] which can be extended without big modifications. Any modification will be enlighted. Some of the new arguments rely on results of [LM08]. As we can factorize the initial condition, we can state this first lemma :

Lemma 4.2.1 (Lemma 3.2.2 of [JJ12]) *For $i \geq 0$ and $\theta \geq 0$, the mappings $x \mapsto x + u_i(\theta, x)$ and $x \mapsto -u_i(\theta, x)$ are non-decreasing, or equivalently, $x \mapsto u_i(\theta, x)$ is 1-Lipschitz and non-increasing.*

By a plain induction on Equation (4.3), for any $(\theta, x) \in \mathbb{R}_+^2$, $(K - x)^+ \leq u_i(\theta, x) \leq K$. The following Lemma needs a proof for the sake of completeness.

Lemma 4.2.2 *The mapping $(\theta, x) \in \mathbb{R}_+^\star \times \mathbb{R}_+ \mapsto u_i(\theta, x) \in \mathbb{R}_+$ is bounded and continuous.*

Proof. We have already mentionned that $0 \leq u_i \leq K$.

We now prove the continuity. We recall that a Lévy process is stochastically continuous and thus for any bounded continuous function f one has trivially $\lim_{t \to 0} \mathbb{E}[f(X_t)] = f(0)$.

Since we have already established Lemma 4.2.1, one just needs to prove the continuity in θ for any $x \geq 0$.

Let $\varepsilon > 0$, by definition of u_i, for any $n \geq 0$, it exists $\tau_\varepsilon^n \in \mathcal{F}_\theta^X$ such that by setting $x_n = n\varepsilon$:

$$u_i(\theta, x_n) - \varepsilon \leq \mathbb{E}\left[e^{-r\tau_\varepsilon^n}\left(\left(K - x_n e^{X_{\tau_\varepsilon^n}}\right)^+ \mathbf{1}_{\{\tau<\theta\}} + \mathbf{1}_{\{\tau_\varepsilon^n=\theta\}} u_i(0, x_n e^{X_\theta})\right)\right] \tag{4.4}$$

And consequently for any $x \geq 0$, by setting $\tau_\varepsilon(x) = \tau_\varepsilon^{\lfloor \frac{x}{\varepsilon} \rfloor}$, we have since $x \mapsto u_i(0,x)$ and $x \mapsto (K-x)$ are non-increasing and 1-Lipschitz and for $x_\varepsilon = \lfloor \frac{x}{\varepsilon} \rfloor \varepsilon$:

$$\begin{aligned}
u_i(\theta, x) - 2\varepsilon &\leq u_i(\theta, x_\varepsilon) - 2\varepsilon \\
&\leq \mathbb{E}\left[e^{-r\tau_\varepsilon(x)}\left(\left(K - x_\varepsilon e^{X_{\tau_\varepsilon(x)}}\right)^+ \mathbf{1}_{\{\tau_\varepsilon(x)<\theta\}} + \mathbf{1}_{\{\tau_\varepsilon(x)=\theta\}} u_i(0, x_\varepsilon e^{X_\theta})\right)\right] - \varepsilon \\
&\leq \mathbb{E}\left[e^{-r\tau_\varepsilon(x)}\left(\left(K - x e^{X_{\tau_\varepsilon(x)}}\right)^+ \mathbf{1}_{\{\tau_\varepsilon(x)<\theta\}} + \mathbf{1}_{\{\tau_\varepsilon(x)=\theta\}} u_i(0, x e^{X_\theta})\right)\right] - \varepsilon \\
&\quad + \mathbb{E}\left[e^{-r\tau_\varepsilon(x)}\left|x - x_\varepsilon\right| e^{X_{\tau_\varepsilon(x)}}\right] \\
&\leq \mathbb{E}\left[e^{-r\tau_\varepsilon(x)}\left(\left(K - x e^{X_{\tau_\varepsilon(x)}}\right)^+ \mathbf{1}_{\{\tau_\varepsilon(x)<\theta\}} + \mathbf{1}_{\{\tau_\varepsilon(x)=\theta\}} u_i(0, x e^{X_\theta})\right)\right]
\end{aligned} \tag{4.5}$$

where we have used the martingale property of $(e^{-rt+X_t})_{t\geq 0}$ and the fact that $|x - x_n| \leq \varepsilon$. Let $\tilde{X}$ be an independent copy of X We define $\tilde{\tau}_\varepsilon(x)$ by analogy with $\tau_\varepsilon(x)$. Then the process $\left(\tilde{\tilde{X}}_s = X_s \mathbf{1}_{\{s\leq t\}} + \tilde{X}_{s-t}\mathbf{1}_{\{s>t\}}\right)_{s\geq 0}$ is again a Lévy process with the same law as X and the following random variable $\tilde{\tilde{\tau}}_\varepsilon = t + \tilde{\tau}_\varepsilon(\bar{S}_t^x)$ is a stopping time of the filtration generated by $\tilde{\tilde{X}}$ and is $\mathbb{P}$-almost surely not greater than $\theta + t$. This stopping time satisfies the property that $\mathbb{P}$-almost surely

$$0 \leq u_i(\theta, xe^{\tilde{\tilde{X}}_t}) - \mathbb{E}\left[e^{-r(\tilde{\tilde{\tau}}_\varepsilon - t)}\left((K - xe^{\tilde{\tilde{X}}_{\tilde{\tilde{\tau}}_\varepsilon}})^+ \mathbf{1}_{\left\{\tilde{\tilde{\tau}}_\varepsilon < \theta+t\right\}} + \mathbf{1}_{\left\{\tilde{\tilde{\tau}}_\varepsilon = \theta+t\right\}} u_i(0, xe^{\tilde{\tilde{X}}_{\theta+t}})\right) |\mathcal{F}_t^X\right] \leq 2\varepsilon.$$

Taking the expectation into each member of the previous inequality gives us that $u_i(\theta+t,x) + 2\varepsilon \geq \mathbb{E}\left[e^{-rt}u_i(\theta, \bar{S}_t^x)\right]$ by definition of $u_i(\theta+t,x)$. As ε is arbitrary, one has $u_i(\theta+t,x) \geq \mathbb{E}\left[e^{-rt}u_i(\theta, \bar{S}_t^x)\right]$. Therefore, using Lemma 4.2.1 for the second inequality, one has :

$$\begin{aligned}
u_i(\theta+t,x) - u_i(\theta,x) \geq& \mathbb{E}\left[e^{-rt}u_i(\theta, \bar{S}_t^x) - u_i(\theta,x)\right] \\
\geq& -rtK - \mathbb{E}\left[\left(\bar{S}_t^x - x\right)^+\right] = -rtK - x\mathbb{E}\left[\left(e^{X_t} - 1\right)^+\right] \\
=& -rtK - x\mathbb{E}\left[e^{X_t} - 1 + \left(1 - e^{X_t}\right)^+\right] \\
\geq& -rtK - x(e^{rt} - 1) - x\mathbb{E}\left[\left(1 - e^{X_t}\right)^+\right]
\end{aligned} \tag{4.6}$$

where we have used the martingale property of $(e^{-rt+X_t})_{t\geq 0}$ for the last line.

By definition of u_i, and since $K, r \geq 0$, for any $\varepsilon > 0$, one has the existence of $\tau_\varepsilon \in \mathcal{F}_{\leq\theta+t}^X$ such that :

$$
\begin{aligned}
u_i(\theta+t,x)-u_i(\theta,x) \leq & \mathbb{E}\left[\left(\left(K-xe^{-r\tau_\varepsilon+X_{\tau_\varepsilon}}\right)^+\mathbf{1}_{\{\tau_\varepsilon<t\}}+\mathbf{1}_{\{\tau_\varepsilon=t\}}u_i(\theta,\bar{S}^x_t)\right)\right]-u_i(\theta,x)+\varepsilon \\
\leq & \mathbb{E}\left[\left(\left(K-xe^{-r\tau_\varepsilon+X_{\tau_\varepsilon}}\right)^+\mathbf{1}_{\{\tau_\varepsilon<t\}}+\mathbf{1}_{\{\tau_\varepsilon=t\}}u_i(\theta,xe^{-rt+X_t})\right)\right] \\
& \qquad -u_i(\theta,x)+\varepsilon \\
\leq & \mathbb{E}\left[u_i(\theta,e^{-rt}\bar{S}^x_t)-u_i(\theta,x)\right]+\varepsilon \\
\leq & \mathbb{E}\left[\left(x-e^{-rt}\bar{S}^x_t\right)^+\right]+\varepsilon=x\mathbb{E}\left[\left(1-e^{-rt+X_t}\right)^+\right]+\varepsilon.
\end{aligned}
\tag{4.7}
$$

where we have used the fact that u_i is non-increasing, that $\left((K-e^{-rt}\bar{S}^x_t)^+\right)_{t\geq 0}$ is a sub-martingale, and that $x\mapsto u_i(\theta,x)\geq (K-x)^+$. Last, we used again Lemma 4.2.1. Since it is true for any $\varepsilon>0$, the inequality stays true with $\varepsilon=0$. Gathering Equations (4.6) and (4.7), we conclude by using the preliminary remark of the proof since $x\mapsto(1-e^x)^+$ is a bounded continuous function and $(X_t)_{t\geq 0}$ and $(-rt+X_t)_{t\geq 0}$ are both Lévy processes. □

Corollary 4.2.3 *Dynamic programming principle holds, i.e for any $0\leq t\leq\theta$ and $x\geq 0$ the following inequality holds for any stopping time $\tau\in\mathcal{F}^X_{\leq\theta}$:*

$$
u_i(\theta,x)\geq\mathbb{E}\left[e^{-r\tau}\left(K-\bar{S}^x_\tau)^+\mathbf{1}_{\{\tau<t\}}+\mathbf{1}_{\{\tau\geq t\}}u_i(\theta-\tau,\bar{S}^x_\tau)\right)\right]
\tag{4.8}
$$

Proof. First apply [PS06, Corollary 2.9 p.46] to get the existence of an optimal stopping time, then apply [PS06, Theorem 2.4 p.37], and endly conclude with Hunt's stopping time theorem ([PS06, p.60]). □

At this point, it is worth to mention that no hypothesis enables us to exclude the case where X is a subordinator. But if we are in this case, X has non-decreasing paths and so for $t,x\geq 0$, since $r\geq 0$ and $x\mapsto(K-x)^+$ is non-increasing and non-negative, $e^{-rt}(K-xe^{X_t})^+\leq e^{-rt}(K-x)^+\leq(K-x)^+$ then it is easy to see that $u_0(\theta,x)=(K-x)^+$. Moreover for the same reasons and by Assumption A_1 on dividend functions, for $i\geq 1$, any optimal stopping time $\tau^\star$ can only take the value 0 or θ and since $\tau^\star\in\mathcal{F}^X_0$, $\mathbb{P}(\tau^\star=0)$ is equal to 0 or 1 and thus the value of $u_i(\theta,x)$ is nothing else than $\max\left((K-x)^+,\mathbb{E}\left[e^{-r\theta}u_{i-1}(\theta^{i-1}_d,\rho_i(xe^{X_\theta}))\right]\right)$.

If $-X$ is a subordinator, then in order to have martingality of e^{-rt+X_t}, we necessarily have $r=0$ and $X\equiv 0$.

We enlight also the case where $r=0$ since in this case, by Jensen's inequality, one gets plainly that $u_0(\theta,x)=\mathbb{E}\left[(K-xe^{X_\theta})^+\right]$ and again by Assumption A_1 and Jensen's inequality since, for $i\geq 1$, $u_{i-1}(\theta^{i-1}_d,\rho_i(y))\geq(K-y)^+$ we get that $u_i(\theta,x)=\mathbb{E}\left[u_{i-1}(\theta^{i-1}_d,\rho_i(xe^{X_\theta}))\right]$.

We then assume that $r>0$, and X is not a subordinator.

But we still need to specify some conditions on X since the previous conditions do not exclude the case where $(X_t-\gamma t)_{t\geq 0}$ is a subordinator for some real number γ. In this case, if the spot price goes too high, there may be no hope to have a positive reward, which contradicts the financial intuition.

For $\theta \geq 0$ and $x \geq Ke^{(-\gamma)^+\theta}$, then for any $t \in [0,\theta]$, $e^{-rt}(K-xe^{X_t})^+ = 0$, and so $u_0(\theta,x) = 0$, and for $i \geq 1$, $u_i(\theta,x) = \mathbb{E}\left[e^{-r\theta}u_{i-1}(\theta_d^{i-1}, \rho_i(xe^{X_\theta})\right]$ which may be null if $u_{i-1}(\theta_d^{i-1}, \rho_i(xe^{\gamma\theta})) = 0$. We decide to exclude this case. And using [CT04, Prop.3.10 p.88] we state that X satisfies the following necessary and sufficient condition.

C_2. At least one of the following conditions is satisfied.

a) $\sigma > 0$,

b) $\int_{(0,1]} |x|\,\nu(dx) = +\infty$,

c) $\nu(\mathbb{R}_-^\star) > 0$.

These conditions (also assumed in [LM08]) guarantee that for any $M \leq 0$ and for any $t > 0$, $\mathbb{P}(X_t \leq M) > 0$. **From now on**, in addition to the assumption on the dividend functions, we assume that $r > 0$ and that C_2 holds.

Corollary 4.2.4 (Corollary 3.2.3 of [JJ12]) *For $i = 0$, u_0 is positive on $\mathbb{R}_+^\star \times \mathbb{R}_+$, and it exists a continous $c_0 : \mathbb{R}_+^\star \to (0,K)$ such that for any $\theta > 0$, $u_0(\theta,x) > (K-x)^+ \Leftrightarrow x > c_0(\theta)$.*

For $i \geq 1$, u_i is positive on $\mathbb{R}_+ \times \mathbb{R}_+$ and for any $\theta \geq 0$, it exists $c_i(\theta) \in [0,K]$ such that $u_i(\theta,x) > (K-x)^+ \Leftrightarrow x > c_i(\theta)$ and c_i is u.s.c.

It is worth to mention that $c_i(\theta)$ *is the so-called exercise boundary of the optimal stopping problem stated in Equation* (4.3). **Proof.** We are going to prove the positivity of u_i by induction on i. Properties of u_0 and c_0 are proved in [LM08], c_0 is the classical exercise boundary of the American put option in the Lévy's exponential model. Let $i \geq 1$, since by induction hypothesis the random variable $u_{i-1}(\theta_d^{i-1}, \rho_i(\bar{S}_\theta^x))$ is positive almost-surely, one has $u_i(\theta,x) \geq \mathbb{E}\left[e^{-r\theta}u_{i-1}(\theta_d^{i-1}, \rho_i(\bar{S}_\theta^x))\right] > 0$. As $u_i(0,x) = u_{i-1}(\theta_d^{i-1}, \rho_i(x))$ and as $\theta_d^0 > 0$, we have the positivity of $u_i(0,.)$.
It remains to prove that $u_i(\theta,x) > (K-x)^+ \Leftrightarrow x > c_i(\theta)$. Let $i \geq 1$, and let us define $c_i(\theta) = \sup\{x \geq 0 | u_i(\theta,x) = (K-x)^+\}$. Since $u_i(\theta,0) = K = (K-0)^+$ and $u_i(\theta,K) > 0$ then $c_i(\theta)$ is well defined and belongs to $[0,K)$. And on $[0,K)$, $u_i(\theta,x) - (K-x)^+ = x + u_i(\theta,x) - (0 + u_i(\theta,0))$, as $x \mapsto x + u_i(\theta,x)$ is non-decreasing, for $x \leq c_i(\theta)$, $u_i(\theta,x) - (K-x)^+ = 0$. Hence it proves the existence of the exercise boundary and at the same time that

$$c_i(\theta) = \inf\{x \geq 0 : u_i(\theta,x) + x > K\} \in \mathbb{R}_+.$$

The upper semi continuity of c_i is then a direct consequence of Lemma 4.2.2. □

The reader familiar to optimal stopping problems can easily see in Equation (4.3) that $(V_t = e^{-rt}u_i(\theta - t, \bar{S}^x_t))_{t\in[0,\theta]}$ is the Snell envelop of the right-continuous process

$$\left(G_t = \max(e^{-rt}(K - \bar{S}^x_t)^+, \mathbb{E}\left[e^{-r\theta}u_i(0, \bar{S}^x_\theta)|\mathcal{F}^X_t\right]\right)_{t\in[0,\theta]}.$$

Indeed, at any time t where $\mathbb{E}\left[e^{-r\theta}u_i(0, \bar{S}^x_\theta)|\mathcal{F}^X_t\right] > e^{-rt}(K - \bar{S}^x_t)^+$, it is always better not to exercise and wait until better days (and maybe until the maturity). For this, it is sufficient to notice that for any $\tau \in \mathcal{F}^X_{\leq\theta}$, by defining $A = \left\{e^{-r\tau}(K - \bar{S}^x_\tau)^+ \geq \mathbb{E}\left[e^{-r\theta}u_i(0, \bar{S}^x_\theta)|\mathcal{F}^X_\tau\right]\right\}$ the following stopping time $\bar{\tau} = \tau\mathbf{1}_{\{A\}} + \theta\mathbf{1}_{\{A^c\}}$ is a better (not necessarily strictly better) strategy. This remark combined with some regularity properties of G stated in Lemma 4.6.1 in Appendix enables us to state the following Proposition which derives directly from [PS06, Th.2.2 p.29].

Proposition 4.2.5 *The stopping time* $\tau^\star = \inf\left\{t \geq 0 : \bar{S}^x_t \leq c_i(\theta - t)\right\} \wedge \theta$ *is optimal in Equation* (4.3).

Proof. A direct application of Theorem 2.2 in Peskir and Shiryaev tells us that the stopping time $\tau = \inf\left\{t \geq 0 : e^{-rt}u_i(\theta - t, \bar{S}^x_t) \leq G_t\right\} \wedge \theta$ is optimal for $u_i(\theta, x)$. Due to the previous remark stating that it is never optimal to exercise when $G_t \neq e^{-rt}(K - \bar{S}^x_t)^+$ except when $t = \theta$, $\tau = \inf\left\{t \geq 0 : u_i(\theta - t, \bar{S}^x_t) \leq (K - \bar{S}^x_t)^+\right\} \wedge \theta$. Due to Corollary 4.2.4, we immediately get that $\tau^\star = \tau$. □

Proposition 4.2.6 (Th.2.4 of [PS06]) *Let* $(\theta, x) \in \mathbb{R}_+ \times \mathbb{R}_+$ *and let* $\tau^\star$ *be the first entry time of the process* $Z^{(\theta,x)}_t \stackrel{\text{def}}{=} (\theta - t, \bar{S}^x_t)$ *into the exercise region :*

$$\mathcal{E}_i = \left\{(\theta', x') \in \mathbb{R} \times \mathbb{R}_+ : \theta' \leq 0 \text{ or } x' \leq c_i(\theta')\right\}.$$

Then the process $e^{-rt\wedge\tau^\star}u_i(Z^{(\theta,x)}_{t\wedge\tau^\star})$ *is a right-continuous martingale.*

4.3 Continuity of the exercise boundary

We will denote by $\bar{c}$ the exercise boundary of the renormalized (i.e with strike 1) American put option in the Lévy's exponential model on $\bar{S}^x$ without dividends. Here are some results of [LM08].

Proposition 4.3.1 (Prop.4.1,Th.4.2 [LM08]) *Under Assumption* C_2*,* $\bar{c}$ *is positive on* $\mathbb{R}_+$*, and is continuous on* $\mathbb{R}^\star_+$*.*

Please note that for any $\theta \geq 0$, $c_0(\theta) = K\bar{c}(\theta)$.

We first state a Lemma which gives a nice comparison between c_i and the exercise boundary of the standard American put option in the Lévy's exponential model. This Lemma enables us first to state that when Assumption C_2 is satisfied, the boundary can not be equal to 0 except at dividends dates. This is a result of high importance to prove regularity results on c_i.

Lemma 4.3.2 (Lemma 3.4.4 of [JJ12]) *For $\theta \geq 0$ and $t \geq 0$, one has :*

$$c_i(\theta + t) \geq \left(K\left(1 - e^{-rt}\right) + c_i(\theta)e^{-rt}\right)\bar{c}(t).$$

The proof of Lemma 4.3.2 is exactly the same as in [JJ12].

Let us define $h = \inf\left\{\delta > 0 : \int_{(-\delta,0)} y^2\nu(dy) > 0\right\}$ with the convention that $\inf\varnothing = +\infty$. First notice that $\nu(\mathbb{R}_-^\star) > 0 \Leftrightarrow h < +\infty$. Consequently, we can split condition C_2 into two parts. Assumption C_2 is then equivalent to either C_{21} or C_{22} is satisfied, where C_{21} and C_{22} are defined below.

C_{21}. At least one of the following conditions is satisfied.

a) $\sigma > 0$,

b) $\int_{(0,1]} |x|\,\nu(dx) = +\infty$,

c) $h = 0$.

C_{22}. Both (a) and (b) are not satisfied and $0 < h < +\infty$.

Using [CT04, Prop.3.10 p.88], we see that C_{22} implies that X is the difference of two subordinators and $0 < h < +\infty$. We first deal with the case C_{21} in the next subsection before treating a particular case of C_{22} in the subsequent subsection.

4.3.1 Under Assumption C_{21}

Proposition 4.3.3 *Under Assumption C_{21}, if c_i is right-continuous on $\mathbb{R}_+^\star$ then for any $\theta > 0$, the mapping $x \mapsto x + u_i(\theta, x)$ is increasing on $(c_i(\theta), +\infty)$ and moreover* $\limsup_{\theta' \uparrow \theta} c_i(\theta') = c_i(\theta)$.

Proof. We are going to exhibit a contradiction if the first statement is not true. Since by Lemma 4.2.1, for any $\theta \geq 0$, $y \mapsto y + u_i(\theta, y)$ is non-decreasing, if the statement is not true then it exists $\theta > 0$ such that $(\theta, y_0) \in \mathcal{E}_i^c$ and $\varepsilon > 0$ such that for any $\eta \in [0, 2\varepsilon]$, $y_0e^\eta + u_i(\theta, y_0e^\eta) = y_0 + u_i(\theta, y_0)$.

Let θ be fixed, and let $\tau^\star$ be the optimal stopping time for (θ, y_0).Then for any $\eta \in [0, 2\varepsilon]$, due to Propositions 4.2.5 and 4.2.6, one has :

$$y_0 e^\eta + u_i(\theta, y_0 e^\eta) = \mathbb{E}\left[e^{-r\tau^\star}\left(\bar{S}^{y_0 e^\eta}_{\tau^\star} + u_i(\theta - \tau^\star, \bar{S}^{y_0 e^\eta}_{\tau^\star})\right)\right] \tag{4.9}$$

and thus :

$$0 = \mathbb{E}\left[e^{-r\tau^\star}\left(\bar{S}^{y_0 e^{2\varepsilon}}_{\tau^\star} + u_i(\theta - \tau^\star, \bar{S}^{y_0 e^{2\varepsilon}}_{\tau^\star}) - \bar{S}^{y_0}_{\tau^\star} - u_i(\theta - \tau^\star, \bar{S}^{y_0}_{\tau^\star})\right)\right]. \tag{4.10}$$

Since by Lemma 4.2.1, for any $\theta' \geq 0$, $y \mapsto y + u_i(\theta', y)$ is non-decreasing, one gets that :

$$\bar{S}^{y_0 e^{2\varepsilon}}_{\tau^\star} + u_i(\theta - \tau^\star, \bar{S}^{y_0 e^{2\varepsilon}}_{\tau^\star}) - \bar{S}^{y_0}_{\tau^\star} - u_i(\theta - \tau^\star, \bar{S}^{y_0}_{\tau^\star}) \overset{\mathbb{P}-\text{a.s}}{=} 0. \tag{4.11}$$

But since the probability that $\tau^\star$ does not coincide with a negative jump of magnitude bigger than ε is greater than the probability there is no negative jump of magnitude bigger than ε on a time interval of length θ, and this last probability is given by $e^{-\theta\nu((-\infty,-\varepsilon))}$, therefore, this event has non-zero probability. Now, Lemma 4.3.2 combined with [LM08, Prop.4.1] implies that the boundary is non-identically zero and then Assumption C_{21} implies that the process X conditionned to have no negative jumps with magnitude bigger than ε satisfies also C_{21} and consequently can hit any finite lower bound on any time interval with a positive probability. Thus, with a positive probability $\tau^\star$ can at the same time be less than θ and not coincide with a negative jump of magnitude bigger than ε. Now, on this event, we have :

$$\forall 0 \leq t < \tau^\star < \theta, \bar{S}^{y_0}_t > c_i(\theta - t), \text{ and } \bar{S}^{y_0}_{\tau^\star} \leq c_i(\theta - \tau^\star) \text{ and } \bar{S}^{y_0}_{\tau^\star} \geq \bar{S}^{y_0}_{\tau^\star_-} e^{-\varepsilon}. \tag{4.12}$$

Now by the right-continuity of c_i, we have :

$$c_i(\theta - \tau^\star) \geq \bar{S}^{y_0}_{\tau^\star} \geq c_i(\theta - \tau^\star) e^{-\varepsilon}. \tag{4.13}$$

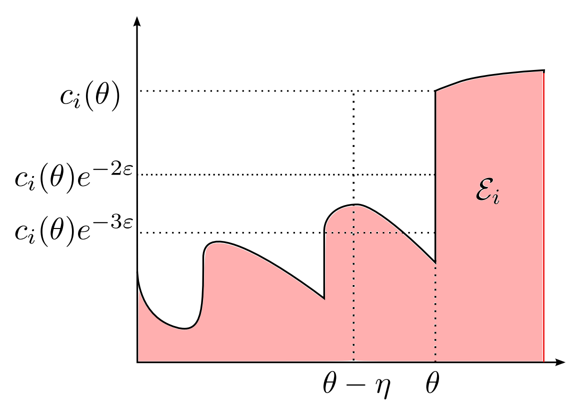
And endly, since $\bar{S}^{y_0 e^{2\varepsilon}}_{\tau^\star} = \bar{S}^{y_0}_{\tau^\star} e^{2\varepsilon} \geq c_i(\theta - \tau^\star) e^{\varepsilon}$ by Equation (4.11), we have :

$$c_i(\theta - \tau^\star) e^\varepsilon + u_i\left(\theta - \tau^\star, c_i(\theta - \tau^\star) e^\varepsilon\right) = c_i(\theta - \tau^\star) + u_i\left(\theta - \tau^\star, c_i(\theta - \tau^\star)\right) = K \tag{4.14}$$

which is the required contradiction. Indeed, this last equality contradicts the definition of the exercise boundary c_i stated in Corollary 4.2.4.

The first part of the statement is proved.

Let us now prove the second part. Again, we will exhibit a contradiction if it is not true. Assume that it is wrong. Then since c_i is u.s.c, it exists $\theta > 0$ and $\varepsilon > 0$ such that $\limsup_{\theta' \uparrow \theta} c_i(\theta') \leq c_i(\theta) e^{-3\varepsilon}$. It then exists η small enough such that for $\theta > \theta' \geq \theta - \eta$ and $x' \geq c_i(\theta) e^{-2\varepsilon}$ we have $(\theta', x') \in \mathcal{E}_i^c$. Configuration is explained on the figure.



By defining $\tau = \inf\left\{t \geq 0 : \bar{S}^1_t \leq e^{-\varepsilon}\right\} \wedge \eta$ and by Proposition 4.2.6, we have for $(\theta', x') \in \mathcal{E}_i^c$ where $\theta' \geq \theta - \eta$, $x' \geq c_i(\theta) e^{-\varepsilon}$:

$$x' + u_i(\theta', x') = \mathbb{E}\left[e^{-r\tau}\left(\bar{S}_\tau^{x'} + u_i(\theta' - \tau, \bar{S}_\tau^{x'})\right)\right]. \tag{4.15}$$

As u_i is bounded and continuous, we can apply the dominated convergence theorem when θ' tends increasingly to θ in the previous expression. It enables us to say that the equality holds with (θ, x) such that $x \in [c_i(\theta)e^{-\varepsilon}, c_i(\theta)]$. Consequently we get :

$$0 = \mathbb{E}\left[e^{-r\tau}\left(\bar{S}_\tau^{c_i(\theta)} + u_i(\theta - \tau, \bar{S}_\tau^{c_i(\theta)}) - \bar{S}_\tau^{c_i(\theta)e^{-\varepsilon}} - u_i(\theta - \tau, \bar{S}_\tau^{c_i(\theta)e^{-\varepsilon}})\right)\right] \tag{4.16}$$

By Lemma 4.2.1, we have in fact that :

$$\bar{S}_\tau^{c_i(\theta)} + u_i(\theta - \tau, \bar{S}_\tau^{c_i(\theta)}) - \bar{S}_\tau^{c_i(\theta)e^{-\varepsilon}} - u_i(\theta - \tau, \bar{S}_\tau^{c_i(\theta)e^{-\varepsilon}}) \overset{\mathbb{P}-\text{a.s}}{=} 0. \tag{4.17}$$

Since $\lim_{\eta\downarrow 0} \mathbb{P}(\inf_{t\in[0,\eta]} \bar{S}_t^1 \geq e^{-\varepsilon}) = 1$ by right-continuity of paths of Lévy processes, for η small enough $\mathbb{P}(\tau = \eta) > 0$ and on this event $(\theta - \eta, \bar{S}_\eta^{c_i(\theta)e^{-\varepsilon}}) \in \mathcal{E}_i^c$ but then the right-hand side of Equation (4.17) should positive by the first part of the statement, which is the required contradiction.

□

Proposition 4.3.4 *Under Assumption* C_{21} *and the hypothesis that* $\int (e^y - 1)^+ \nu(dy) \leq r$*, the function* $\theta \geq 0 \mapsto c_i(\theta)$ *is continuous.*

Proof. Since under the assumption that $\int (e^y - 1)^+ \nu(dy) \leq r$, by [LM08, Th.4.4] $\lim_{t\downarrow 0} \bar{c}(t) = 1$, Lemma 4.3.2 implies that $\liminf_{\theta'\downarrow\theta} c_i(\theta') \geq c_i(\theta)$. The upper-semicontinuity property stated in Lemma 4.2.1 implies that c_i is right-continuous and then we just need to prove that it does not exist $\theta > 0$ such that $\liminf_{t\downarrow 0} c_i(\theta - t) < c_i(\theta)$.

Let $c_- \overset{\text{def}}{=} \liminf_{t\downarrow 0} c_i(\theta - t)$ and $(t_n)_n$ be a decreasing sequence in $(0, \theta)$ tending to zero and such that $c_i(\theta - t_n)$ tend to c_-. Then, by Lemma 4.3.2 written with $(s - t_n, \theta - s)$ replacing (t, θ), we obtain that for $s \in (t_n, \theta)$, $c_i(\theta - s) \leq c_i(\theta - t_n)\dfrac{e^{r(s-t_n)}}{\bar{c}(s - t_n)}$. So $\lim_{t\downarrow 0} c_i(\theta - t) = c_-$.

We have shown that $\liminf_{t\downarrow 0} c_i(\theta - t) = \limsup_{t\downarrow 0} c_i(\theta - t)$, and then we conclude by using Proposition 4.3.3. □

4.3.2 An interesting particular case

We first state a Lemma which is valid under Assumption C_2.

Lemma 4.3.5 *Let* $\mathfrak{e}_r$ *be an exponential random variable independent of* $\mathcal{F}^X$*, then for* $\theta, x \geq 0$ *and* $0 \leq t \leq \theta$ *one has :*

$$x + u_i(\theta, x) = \sup_{\tau\in\mathcal{F}^X} \mathbb{E}\left[\mathbf{1}_{\{\tau<\mathfrak{e}_r\}}\left(K\mathbf{1}_{\{\tau<t\}} + (\bar{S}_t^x + u_i(\theta - t, \bar{S}_t^x))\mathbf{1}_{\{\tau\geq t\}}\right)\right]. \tag{4.18}$$

Proof. Let $\theta, x \geq 0$. Since $(e^{-rt}\bar{S}_t^x)_{t\geq 0}$ is a martingale and by the dynamic programming principle stated in Corollary 4.2.3, for $0 \leq t \leq \theta$:

$$x + u_i(\theta, x) = \sup_{\tau \in \mathcal{F}^X_{\leq \theta}} \mathbb{E}\left[e^{-r\tau}\left(\bar{S}_\tau^x + (K - \bar{S}_\tau^x)^+ \mathbf{1}_{\{\tau < t\}} + u_i(\theta - t, \bar{S}_t^x)\mathbf{1}_{\{\tau \geq t\}}\right)\right]. \tag{4.19}$$

With the notations of Proposition 4.2.5, and since $c_i(\theta) \leq K$ for $\theta \in \mathbb{R}_+$ we have :

$$\begin{aligned} x + u_i(\theta, x) =& \mathbb{E}\left[e^{-r\tau^\star}\left(K\mathbf{1}_{\{\tau^\star < t\}} + (\bar{S}_t^x + u_i(\theta - t, \bar{S}_t^x))\mathbf{1}_{\{\tau^\star \geq t\}}\right)\right] \\ &\leq \sup_{\tau \in \mathcal{F}^X} \mathbb{E}\left[e^{-r\tau}\left(K\mathbf{1}_{\{\tau < t\}} + (\bar{S}_t^x + u_i(\theta - t, \bar{S}_t^x))\mathbf{1}_{\{\tau \geq t\}}\right)\right] \\ &\leq \sup_{\tau \in \mathcal{F}^X_{\leq \theta}} \mathbb{E}\left[e^{-r\tau}\left(\bar{S}_\tau^x + (K - \bar{S}_\tau^x)^+\mathbf{1}_{\{\tau < t\}} + u_i(\theta - t, \bar{S}_t^x)\mathbf{1}_{\{\tau \geq t\}}\right)\right] \\ &= x + u_i(\theta, x) \end{aligned} \tag{4.20}$$

Now, since $\mathfrak{e}_r$ is independent with respect to $\mathcal{F}^X$, we have $\mathbb{E}\left[\mathbf{1}_{\{\tau < \mathfrak{e}_r\}} | \mathcal{F}^X\right] = e^{-r\tau}$ $\mathbb{P}$-almost surely, and this proves the statement. $\square$

Now, we assume in addition of the previous assumptions that it exists $\mathrm{d} > 0$ such that $(\mathrm{d}t - X_t)_{t \geq 0}$ is both a subordinator with no drift and a compound Poisson process.

Lemma 4.3.6 *Assume that it exists an increasing function ϕ_i from $\mathbb{R}_+$ to $\mathbb{R}_+$ null at zero such that for any $y \geq x \geq c_i(0)$, $y + u_i(0, y) - (x + u_i(0, x)) \geq \phi_i(y - x)$. Then for $\theta', \theta > 0$ such that $c_i(\theta) \geq c_i(\theta')$, one has :*

$$\left|c_i(\theta) - c_i(\theta')\right| \leq e^{-\mathrm{d}\theta'}\phi_i^{(-1)}\left(Ke^{(r + \nu(\mathbb{R}_-))\theta'}\left(2(e^{r|\theta - \theta'|} - 1) + 1 - e^{-(\mathrm{d} - r)|\theta - \theta'|}\right)\right) \tag{4.21}$$

where $\phi_i^{(-1)}$ is the inverse function of ϕ_i.

This result proves the existence of a uniform modulus of continuity on any compact for the exercise boundary. Notice that this Lemma is true for $i = 0$ with $\phi_0(x) = x$. In Lemma 4.3.7, we will provide sufficient conditions on the dividend functions to get the existence of such functions. **Proof.** By the previous Lemma, for any $\theta, x \geq 0$ and $0 \leq t \leq \theta$, we obtain that :

$$x + u_i(\theta, x) \leq \mathbb{E}\left[\sup_{0 \leq s \leq \mathfrak{e}_r}\left(K\mathbf{1}_{\{s < t\}} + (\bar{S}_t^x + u_i(\theta - t, \bar{S}_t^x))\mathbf{1}_{\{s \geq t\}}\right)\right] \tag{4.22}$$

Since $x + u_i(\theta - t, x) \geq K$, $s \mapsto K\mathbf{1}_{\{s < t\}} + \left(\bar{S}_t^x + u_i(\theta - t, \bar{S}_t^x)\right)\mathbf{1}_{\{s \geq t\}}$ is non-decreasing. As $\mathfrak{e}_r$ is independent of $\mathcal{F}^X$ we get :

$$\begin{aligned} \mathbb{E}\left[\sup_{0 \leq s \leq \mathfrak{e}_r}\left(K\mathbf{1}_{\{s < t\}} + (\bar{S}_t^x + u_i(\theta - t, \bar{S}_t^x))\mathbf{1}_{\{s \geq t\}}\right)\right] &= \mathbb{E}\left[K\mathbf{1}_{\{\mathfrak{e}_r < t\}} + (\bar{S}_t^x + u_i(\theta - t, \bar{S}_t^x))\mathbf{1}_{\{\mathfrak{e}_r \geq t\}}\right] \\ &= \mathbb{E}\left[K(1 - e^{-rt}) + e^{-rt}(\bar{S}_t^x + u_i(\theta - t, \bar{S}_t^x))\right] \end{aligned} \tag{4.23}$$

Since $K = K(1-e^{-rt}) + e^{-rt}(\bar{S}_t^0 + u_i(\theta - t, \bar{S}_t^0))$, the following quantity is well defined on $[0, +\infty]$:

$$\gamma_i(\theta, t) \stackrel{\text{def}}{=} \sup\left\{x \geq 0 : K = K(1-e^{-rt}) + x + e^{-rt}\mathbb{E}\left[u_i(\theta - t, \bar{S}_t^x)\right]\right\} \tag{4.24}$$

Since for $x \leq \gamma_i(\theta, t)$, one has $K \geq x + u_i(\theta, x)$, one gets that $c_i(\theta) \geq \gamma_i(\theta, t)$.

But in the particular case of this section, $\gamma_i(\theta, t) = c_i(\theta - t)e^{-\mathrm{d}t}$. Indeed, for $x \leq c_i(\theta - t)e^{-\mathrm{d}t}$, one has $\mathbb{P}(\bar{S}_t^x \leq c_i(\theta - t)) = 1$ which implies $\mathbb{E}\left[e^{-rt}u_i(\theta - t, \bar{S}_t^x)\right] = Ke^{-rt} - x$ and therefore $\gamma_i(\theta, t) \geq c_i(\theta - t)e^{-\mathrm{d}t}$. Now since $\mathbb{P}(X_t = \mathrm{d}t) = e^{-t\nu(\mathbb{R}_-)} > 0$, for $x > c_i(\theta - t)e^{-\mathrm{d}t}$, $\mathbb{P}(\bar{S}_t^x > c_i(\theta - t)) \geq \mathbb{P}(X_t = \mathrm{d}t) > 0$.

We then get that for $0 \leq t \leq \theta$

$$\gamma_i(\theta, t) = c_i(\theta - t)e^{-\mathrm{d}t} \leq c_i(\theta). \tag{4.25}$$

The first jump time of X is denoted by T and the jump size by ΔX. For sake of simplicity, we introduce for any $t, y \geq 0$, $v_i(t, y) = y + u_i(t, y)$.

Let $\theta, x \geq 0$ such that $x > c_i(\theta)$, since for $t \leq T$, $X_t = \mathrm{d}t$ and $c_i(\theta - t)e^{-\mathrm{d}t} \leq c_i(\theta)$, we get that $\tau^\star \geq T \wedge \theta$. By Proposition 4.2.6, one has :

$$\begin{aligned} v_i(\theta, x) =& \mathbb{E}\left[e^{-rT}v_i(\theta - T, xe^{\mathrm{d}T - \Delta X})\mathbf{1}_{\{T \leq \theta\}} + e^{-r\theta}v_i(0, xe^{\mathrm{d}\theta})\mathbf{1}_{\{T > \theta\}}\right]. \\ =& x + \mathbb{E}\left[e^{-rT}u_i(\theta - T, xe^{\mathrm{d}T - \Delta X})\mathbf{1}_{\{T \leq \theta\}} + e^{-r\theta}u_i(0, xe^{\mathrm{d}\theta})\mathbf{1}_{\{T > \theta\}}\right]. \end{aligned} \tag{4.26}$$

Since u_i is bounded and continuous, using the dominated convergence theorem, we get that Equation (4.26) is still valid with $x = c_i(\theta)$.

For $y \geq x \geq c_i(\theta)$, using Equation (4.26) and the monotonicity of v_i in the state variable, one gets :

$$v_i(\theta, y) - v_i(\theta, x) \geq e^{-(r + \nu(\mathbb{R}_-))\theta}\left(v_i(0, ye^{\mathrm{d}\theta}) - v_i(0, xe^{\mathrm{d}\theta})\right). \tag{4.27}$$

Let $\theta', \theta > 0$ and $c_i(\theta) \geq c_i(\theta')$, we get :

$$\begin{aligned} u_i(\theta', c_i(\theta)) - u_i(\theta, c_i(\theta)) &= c_i(\theta) + u_i(\theta', c_i(\theta)) - (c_i(\theta') + u_i(\theta', c_i(\theta'))) \\ &= v_i(\theta', c_i(\theta)) - v_i(\theta', c_i(\theta')) \\ &\geq e^{-(r + \nu(\mathbb{R}_-))\theta'}\left(v_i(0, c_i(\theta)e^{\mathrm{d}\theta'}) - v_i(0, c_i(\theta')e^{\mathrm{d}\theta'})\right) \\ &\geq e^{-(r + \nu(\mathbb{R}_-))\theta'}\phi_i\left(\left(c_i(\theta) - c_i(\theta')\right)e^{\mathrm{d}\theta'}\right) \end{aligned} \tag{4.28}$$

where we have used the assumption of the statement and the fact that $c_i(\theta') \geq c_i(0)e^{-\mathrm{d}\theta'}$. Endly it gives us :

$$0 \leq c_i(\theta) - c_i(\theta') \leq e^{-\mathrm{d}\theta'}\phi_i^{(-1)}\left(e^{(r + \nu(\mathbb{R}_-))\theta'}\left(u_i(\theta', c_i(\theta)) - u_i(\theta, c_i(\theta))\right)\right) \tag{4.29}$$

and noticing that for any $\theta \geq 0$, $c_i(\theta) \leq K$ we get the following Equation by using Equations (4.6) and (4.7) in the proof of Lemma 4.2.2.

$$\left|c_i(\theta) - c_i(\theta')\right| \le e^{-\mathrm{d}\theta'} \phi_i^{(-1)} \left(K e^{(r+\nu(\mathbb{R}_-))\theta'} \left(2(e^{r|\theta-\theta'|}-1) + \mathbb{E}\left[(1-e^{-r|\theta-\theta'|+X_{|\theta-\theta'|}})^+\right]\right)\right)$$

We then notice that since $\psi(1) = r$, and since $(\mathrm{d}t - X_t)_{t\ge 0}$ is a subordinator with no drift so for $t \ge 0$, $1 \le \mathbb{E}\left[e^{-(rmdt-X_t)}\right] = e^{(r-\mathrm{d})t}$ and we get that $\mathrm{d} \ge r$. Consequently one gets for any $t \ge 0$:

$$\mathbb{E}\left[(1-e^{-rt+X_t})^+\right] \le \mathbb{E}\left[(1-e^{-\mathrm{d}t+X_t})^+\right] \le 1 - \mathbb{E}\left[e^{-\mathrm{d}t+X_t}\right] = 1 - e^{-(\mathrm{d}-r)t}$$

□

Lemma 4.3.7 *If for any $j \le i$, we define $d_j^\star = \inf\{x \ge 0 : D_j(x) > 0\}$ and we assume that it exists two increasing function $\overline{\delta}_j$ and $\underline{\delta}_j$ from $\mathbb{R}_+$ to $\mathbb{R}_+$ such that for $y \ge x \ge d_j^\star$, $D_j(y) - D_j(x) \ge \underline{\delta}_j(y-x)$ and for $y \ge x \ge 0$ $y - D_j(y) - (x - D_j(x)) \ge \overline{\delta}_j(y-x)$. Then by setting $\phi_0(z) = z$ and inductively for $j < i$:*

$$\phi_{j+1}(z) = \begin{cases} e^{-(r+\nu(\mathbb{R}_-))\theta_d^j} \phi_j(e^{\mathrm{d}\theta_d^j} \overline{\delta}_{j+1}(z)) \text{ if } d_{j+1}^\star > c_j(\theta_d^j) \\ \underline{\delta}_{j+1}(z) \text{ if } d_{j+1}^\star \le c_j(\theta_d^j) \end{cases}. \tag{4.30}$$

One gets that for $j \le i$, for $y \ge x \ge c_j(0)$, $y + u_j(0,y) - (x + u_j(0,x)) \ge \phi_j(y-x)$ and ϕ_j is increasing and null at zero.

With Lemma 4.3.7, we deduce the following Corollary.

Corollary 4.3.8 *With the notations of the previous Lemma, assume that it exists two reals numbers α, β such that $0 < \alpha \le \beta < 1$ and for any $j \le i$, for $y \ge x \ge d_i^\star$, $\alpha(y-x) \le D_i(y) - D_i(x)$ and for $y \ge x \ge 0$, $D_i(y) - D_i(x) \le \beta(y-x)$ then for any $1 \le j \le i$, for any $\theta, \theta' \ge 0$ such that $c_j(\theta) \ge c_j(\theta')$ one has :*

$$\begin{aligned} 0 \le c_j(\theta) - c_j(\theta') \le e^{-\mathrm{d}\theta'} \min(\alpha, (1-\beta))^{-j} e^{-(\mathrm{d}-(r+\nu(\mathbb{R}_-))\sum_{k=0}^{j-1}\theta_d^k} \\ \times \left(K e^{(r+\nu(\mathbb{R}_-))\theta'} \left(2(e^{r|\theta-\theta'|}-1) + 1 - e^{-\mathrm{d}|\theta-\theta'|}\right)\right). \end{aligned} \tag{4.31}$$

Proof. First we notice that since $\psi(1) = r \ge \mathrm{d} - \nu(\mathbb{R}_-)$, we get that $\mathrm{d} - (r + \nu(\mathbb{R}_-)) \le 0$. Therefore by Lemma 4.3.7, we get for any $1 \le j \le i$ and for any $y \ge x \ge c_j(0)$:

$$y + u_j(0,y) - (x + u_j(0,x)) \ge \min(\alpha, (1-\beta))^j e^{(\mathrm{d}-(r+\nu(\mathbb{R}_-))k=0^{j-1}\theta_d^k}(y-x). \tag{4.32}$$

We then apply Lemma 4.3.6 to each j and we get the result. □

We now prove Lemma 4.3.7. **Proof.** For $i = 0$, since $u_0(0,x) = (K-x)^+$ and $c_0(0) = K$, one gets that for $y \ge x \ge K$, $y + u_0(0,y) - (x + u_0(0,x)) = y - x$. We are going to prove the statement by induction. Let $j \ge 1$, and let us assume that at rank $j-1$, we have for $y \ge x \ge c_{j-1}(0)$, $y + u_{j-1}(0,y) - (x + u_{j-1}(0,x)) \ge \phi_{j-1}(y-x)$. For an easier reading, let us recall that ρ_j is the function which maps $x \ge 0$ into $x - D_j(x) \ge 0$. For $y \ge x \ge 0$, we have :

$$\begin{aligned}\varDelta &:= y + u_j(0,y) - (x + u_j(0,x)) \\ &= D_j(y) - D_j(x) \\ &\quad + \rho_j(y) + u_{j-1}(\theta_d^{j-1}, \rho_j(y)) - \Big(\rho_j(x) + u_{j-1}(\theta_d^{j-1}, \rho_j(x))\Big)\,.\end{aligned} \tag{4.33}$$

Plugging Equation (4.27) into Equation (4.33), we get :

$$\begin{aligned}\varDelta &\geq D_j(y) - D_j(x) \\ &\quad + e^{-(r+\nu(\mathbb{R}_-))\theta_d^{j-1}} \begin{pmatrix} \rho_j(y)e^{\mathrm{d}\theta_d^{j-1}} + u_{j-1}(0, \rho_j(y)e^{\mathrm{d}\theta_d^{j-1}}) \\ -\rho_j(x)e^{\mathrm{d}\theta_d^{j-1}} - u_{j-1}(0, \rho_j(x)e^{\mathrm{d}\theta_d^{j-1}}) \end{pmatrix} \mathbf{1}_{\{\rho_j(x)\geq c_{j-1}(\theta_d^{j-1})\}}.\end{aligned} \tag{4.34}$$

By Equation (4.25), we can apply the induction hypothesis at rank $j-1$ into Equation (4.34) to get :

$$\begin{aligned}\varDelta &\geq D_j(y) - D_j(x) \\ &\quad + e^{-(r+\nu(\mathbb{R}_-))\theta_d^{j-1}} \phi_{j-1}\left(e^{\mathrm{d}\theta_d^{j-1}}\left(\rho_j(y) - \rho_j(x)\right)\right) \mathbf{1}_{\{x - D_j(x) \geq c_{j-1}(\theta_d^{j-1})\}} \\ &\geq \underbrace{\underline{\delta}_j(y-x)\mathbf{1}_{\{x\geq d_j^\star\}}}_{(a)} \\ &\quad + \underbrace{e^{-(r+\nu(\mathbb{R}_-))\theta_d^{j-1}} \phi_{j-1}\left(e^{\mathrm{d}\theta_d^{j-1}}\overline{\delta}_j\,(y-x)\right) \mathbf{1}_{\{x - D_j(x) \geq c_{j-1}(\theta_d^{j-1})\}}}_{(b)}\end{aligned} \tag{4.35}$$

where for the last inequality, we have used growth assumptions on D_j and ρ_j. Both terms (a) and (b) are non-negative.

Since for any $x \geq 0$, $u_j(0,x) = u_{j-1}(\theta_d^{j-1}, x - D_j(x))$, we have $c_j(0) = \min(d_j^\star, c_{j-1}(\theta_d^{j-1}))$. If $c_j(0) = d_j^\star$ then the recursive construction of the statement holds at rank j by using the lower bound (a). If $c_j(0) = c_{j-1}(\theta_d^{j-1})$ then it holds by using the lower bound (b). We then have build an increasing ϕ_j null at zero such that the induction hypothesis holds at rank j. □

4.4 Toward smooth fit

We mention a Lemma which can be used as a first step in a proof of the smooth-fit property, as it is suggested in the conclusion of [AK05].

Lemma 4.4.1 (Lemma 3.4.9 of [JJ12]) *Assume that for some $t_0 > 0$, X_{t_0} has a density with respect to the Lebesgue measure. For any $x \geq 0$, the right-hand derivative of $u_i(\theta, .)$ is well defined and is denoted by $\partial_x u_i(\theta, x)$. Let $\theta > 0$, $x \geq 0$ and τ be an optimal stopping time for $u_i(\theta, x)$. Then one has*

$$1 + \partial_x u_i(\theta, x) \geq \mathbb{E}^{\mathbb{Q}}\left[\mathbf{1}_{\{\tau=\theta\}}\left(1 + \partial_x u_i(0, \bar{S}_\theta^x)\right)\right].$$

Moreover, $\overline{\tau} \overset{\text{def}}{=} \lim_{\epsilon\to 0_+} \inf\left\{t \geq 0 | \bar{S}_t^{x+\epsilon} \leq c_i(\theta - t)\right\}$ *is an optimal stopping time for* $u_i(\theta, x)$ *and satisfies*

$$1 + \partial_x u_i(\theta, x) = \mathbb{E}^{\mathbb{Q}}\left[\mathbf{1}_{\{\overline{\tau}=\theta\}}\left(1 + \partial_x u_i(0, \bar{S}_\theta^x)\right)\right] = \mathbb{E}\left[\frac{e^{-r\theta}\bar{S}_\theta^x}{x}\mathbf{1}_{\{\overline{\tau}=\theta\}}\left(1 + \partial_x u_i(0, \bar{S}_\theta^x)\right)\right].$$

Proof. Since by assumption, for some $t_0 > 0$, X_{t_0} has a density with respect to the Lebesgue measure, for any $t > 0$ X_t has a density with respect to the Lebesgue measure as it is noticed in Exercise 4 of [Ber96, Ch.I §5 p.39]. Thus it enables us to follow exactly the same proof as the one of Lemma 3.4.9 of [JJ12]. □

Assume that the assumptions of Lemma 4.4.1 are fulfilled, and let us introduce the bi-dimensional Lévy process $Y = (Y^{(1)}, Y^{(2)})$ defined by its characteristic exponent $\tilde{\psi} : \lambda_1 \times \lambda_2 \in \mathbb{C} \times \Lambda \mapsto -\lambda_1 + \psi(\lambda_2)$ and let $\mathbb{P}_{(\theta,x)}$ stand for the probability under which Y starts from $(\theta, \ln x)$. Let us define $B = \left\{(t, x) \in \mathbb{R}^2 : t \leq 0 \text{ or } e^x \leq c_i(t)\right\}$ and $T_B = \inf\{t \geq 0 : Y_t \in B\}$. Then setting $(X_t = Y_t^{(2)} - Y_0^{(2)})_{t\geq 0}$ for $\theta, x \geq 0$, the optimal stopping time $\tau^\star$ for (θ, x) defined in Proposition 4.2.5 is $\mathbb{P}_{(\theta,x)}$-almost surely equal to T_B. We then can rewrite the result of Lemma 4.4.1 :

$$1 + \partial_x u_i(\theta, x) = \lim_{\varepsilon\downarrow 0} \mathbb{E}_{(\theta,x+\varepsilon)}\left[\frac{e^{-r\theta+Y_\theta^{(2)}}}{x+\varepsilon}\mathbf{1}_{\{T_B\geq\theta\}}\left(1 + \partial_x u_i(0, e^{Y_\theta^{(2)}})\right)\right]. \tag{4.36}$$

Consequently, by the dominated convergence theorem, we will have the smooth-fit property if and only if $\lim_{\varepsilon\downarrow 0} \mathbb{P}_{(\theta,x+\varepsilon)}(T_B \geq \theta) = 0$.

4.5 Approximation of the value function

We will denote by $\bar{u}$ the value function of the renormalized (i.e with strike 1) American put option in the Lévy's exponential model on $\bar{S}^x$ without dividends. We introduce for $\delta > 0$, the function w_i^δ defined for $\theta, x \geq 0$ by :

$$w_i^\delta(\theta, x) = \begin{cases} \max\left(\bar{u}(\theta, x), \mathbb{E}\left[e^{-r\theta}\dfrac{u_i(0, K\bar{S}_\theta^x)}{K}\right]\right) & \text{if } \theta \in [0, \delta] \text{ (a)} \\ \max\left(\bar{u}(\delta, x), \mathbb{E}\left[e^{-r\delta}w_i^\delta(\theta - \delta, \bar{S}_\delta^x)\right]\right) & \text{if } \theta \geq \delta \quad \text{(b)} \end{cases} \tag{4.37}$$

Firstly, by Equation (4.37.a), w_i^δ is well defined for $(\theta, x) \in [0, \delta] \times \mathbb{R}_+$ and by Equation (4.37.b), w_i^δ is then defined for $(\theta, x) \in \cup_{k\in\mathbb{N}}[k\delta, (k+1)\delta] \times \mathbb{R}_+$ by induction on k. Secondly, we remark that for any $x \geq 0$, $w_i^\delta(0, x) = \frac{u_i(0,Kx)}{K}$ because $\bar{u}(0, x) = (1 - x)^+ \leq \frac{u_i(0,Kx)}{K}$, thus for $\theta = \delta$ both Equations (4.37.a) and (4.37.b) define the same function $x \geq 0 \mapsto w_i^\delta(\delta, x)$.

Proposition 4.5.1 *Let $\delta > 0$ and let w_i^δ be defined by Equation* (4.37)*. Then we have for any $\theta, x \geq 0$.*

$$K w_i^\delta\left(\theta, \frac{x}{K}\right) \leq u_i(\theta, x) \leq K e^{r\delta} w_i^\delta\left(\theta, \frac{x}{K e^{r\delta}}\right) \tag{4.38}$$

The same result holds when for any $t, y \geq 0$, $\bar{u}(t, y)$ is replaced by the smaller quantity $\mathbb{E}\left[e^{-rt}(1 - \bar{S}_t^y)^+\right]$ in the definition of w_i^δ. It comes from the first inequality in Equation (4.44) stated in the next lines. **Proof.** Let $\theta \geq 0$ and let $\mathcal{G}$ be the discrete-time filtration generated by the discrete-time process $\left((\bar{S}^x_{\min(n\delta,\theta)})_{n\geq 0}\right)$. The notation $\tau \in \mathcal{G}$ means that τ is a stopping time of the filtration $\mathcal{G}$. Then for $x \geq 0$, we have that

$$w_i^\delta(\theta', x) = \sup_{\tau \in \mathcal{G}} \mathbb{E}\left[e^{-r\tau}\left(\bar{u}(\tau, \bar{S}_\tau^x)\mathbf{1}_{\{\tau < \theta'\}} + \mathbf{1}_{\{\tau \geq \theta'\}} \frac{u_i(0, K\bar{S}^x_{\theta'})}{K}\right)\right]. \tag{4.39}$$

Indeed, it is well known (cf.[PS06, Chap.I]) that the discrete time optimal stopping problem (4.39) is the unique solution of the Wald-Bellman equation (4.37). Moreover it exists a stopping time $\tau^\star$ of $\mathcal{G}$ taking values on $\left\{k\delta : k \in \mathbb{N}, 0 \leq k \leq \lfloor\frac{\theta}{\delta}\rfloor\right\} \cup \{\theta\}$ such that :

$$w_i^\delta(\theta, x) = \mathbb{E}\left[e^{-r\tau^\star}\left(\bar{u}(\tau^\star, \bar{S}^x_{\tau^\star})\mathbf{1}_{\{\tau^\star < \theta\}} + \mathbf{1}_{\{\tau^\star \geq \theta\}} \frac{u_i(0, K\bar{S}^x_\theta)}{K}\right)\right]. \tag{4.40}$$

Let us define for $t, y \geq 0$, $\bar{\tau}(t, y, (X_s)_{s\in[0,t]})$ as the optimal stopping time rule for the standard American put option in the Lévy's exponential model when the maturity is t, the initial spot price process is y and the dynamics of the log-price process is driven by X. We introduce $\bar{\tau}^\star$ defined by :

$$\bar{\tau}^\star = \left(\begin{array}{l} \sum_{n=0}^{\lfloor\frac{\theta}{\delta}\rfloor - 1}\left(\bar{\tau}\left(\delta, \bar{S}^{\frac{x}{K}}_{n\delta}, (X_{n\delta+s} - X_{n\delta})_{s\in[0,\delta]}\right) + n\delta\right)\mathbf{1}_{\{\tau^\star = n\delta\}} \\ + \left(\bar{\tau}\left(\theta - \lfloor\frac{\theta}{\delta}\rfloor\delta, \bar{S}^{\frac{x}{K}}_{\lfloor\frac{\theta}{\delta}\rfloor\delta}, (X_{\lfloor\frac{\theta}{\delta}\rfloor\delta+s} - X_{\lfloor\frac{\theta}{\delta}\rfloor\delta})_{s\in[0,\theta-\lfloor\frac{\theta}{\delta}\rfloor\delta]}\right) + \lfloor\frac{\theta}{\delta}\rfloor\delta\right)\mathbf{1}_{\{\tau^\star = \lfloor\frac{\theta}{\delta}\rfloor\delta < \theta\}} \\ + \theta\mathbf{1}_{\{\tau^\star = \theta\}} \end{array}\right).$$

And we get that :

$$\begin{aligned} a :=& \mathbb{E}\left[e^{-r\bar{\tau}^\star}\left((K - \bar{S}^x_{\bar{\tau}^\star})^+\mathbf{1}_{\{\bar{\tau}^\star < \theta\}} + \mathbf{1}_{\{\bar{\tau}^\star = \theta\}} u_i(0, \bar{S}^x_\theta)\right)\right] \\ =& \mathbb{E}\left[\mathbb{E}\left[e^{-r\bar{\tau}^\star}\left((K - \bar{S}^x_{\bar{\tau}^\star})^+\mathbf{1}_{\{\bar{\tau}^\star < \theta\}} + \mathbf{1}_{\{\bar{\tau}^\star = \theta\}} u_i(0, \bar{S}^x_\theta)\right)\Big| \mathcal{G}\right]\right] \\ =& \mathbb{E}\left[\mathbb{E}\left[e^{-r\tau^\star}\left(K\bar{u}(\tau^\star, \bar{S}^{\frac{x}{K}}_{\tau^\star})\mathbf{1}_{\{\tau^\star < \theta\}} + \mathbf{1}_{\{\tau^\star \geq \theta\}} u_i(0, K\bar{S}^{\frac{x}{K}}_\theta)\right)\right]\right] = K w_i^\delta\left(\theta, \frac{x}{K}\right) \end{aligned} \tag{4.41}$$

using the factorization property of the initial condition which states that $\left(\bar{S}_t^{xK}\right)_{t\geq 0} = \left(K\bar{S}_t^x\right)_{t\geq 0}$ and the independence of the increments of X. It proves the left hand inequality of Equation (4.38).

We now prove the right hand inequality. Using again the factorization property of the initial condition which states that $\left(\bar{S}_t^{xKe^{r\delta}}\right)_{t\geq 0} = \left(Ke^{r\delta}\bar{S}_t^x\right)_{t\geq 0}$,

Then by the same argument stated to justify Equation (4.7), we get for $\theta \leq \delta$:

$$
\begin{aligned}
u_i(\theta, Ke^{r\delta}x) &\leq \max\left(K\mathbb{E}\left[(1-\bar{S}^x_\theta)^+\right], e^{-r\theta}\mathbb{E}\left[u_i(0, Ke^{r\delta}\bar{S}^x_\theta)\right]\right) \\
&\leq \max\left(Ke^{r\theta}\bar{u}(\theta, x), e^{-r\theta}\mathbb{E}\left[u_i(0, Ke^{r\delta}\bar{S}^x_\theta)\right]\right) \\
&\leq Ke^{r\delta}\max\left(\bar{u}(\theta, x), \frac{e^{-r\theta}\mathbb{E}\left[u_i(0, K\bar{S}^x_\theta)\right]}{K}\right) = Ke^{r\delta}w_i^\delta(\theta, x)
\end{aligned}
\tag{4.42}
$$

where we used the monotonicity and the non-negativity of u_i for the last inequality. We have proved that for $\theta \leq \delta$, the right hand inequality of Equation (4.38) holds. Let us assume that for some $n \geq 1$, it holds for $\theta \leq n\delta$. By applying the dynamic programming principle at δ, we get for $\theta \in [n\delta, (n+1)\delta]$ and $x \geq 0$:

$$
u_i(\theta, Ke^{r\delta}x) = \sup_{\tau \in \mathcal{F}^X_{\leq\theta}} \mathbb{E}\left[e^{-r\tau}(K - Ke^{r\delta}\bar{S}^x_\tau)^+\mathbf{1}_{\{\tau<\delta\}} + \mathbf{1}_{\{\tau\geq\delta\}}e^{-r\delta}u_i(\theta-\delta, Ke^{r\delta}\bar{S}^x_\delta)\right]. \tag{4.43}
$$

Using again the same argument stated to justify Equation (4.7), one deduces :

$$
\begin{aligned}
u_i(\theta, Ke^{r\delta}x) &\leq \max\left(K\mathbb{E}\left[(1-\bar{S}^x_\delta)^+\right], e^{-r\delta}\mathbb{E}\left[u_i(\theta-\delta, Ke^{r\delta}\bar{S}^x_\delta)\right]\right) \\
&\leq \max\left(Ke^{r\delta}\bar{u}(\delta, x), e^{-r\delta}\mathbb{E}\left[u_i(\theta-\delta, Ke^{r\delta}\bar{S}^x_\delta)\right]\right) \\
&\leq \max\left(Ke^{r\delta}\bar{u}(\delta, x), e^{-r\delta}\mathbb{E}\left[Ke^{r\delta}w_i(\theta-\delta, \bar{S}^x_\delta)\right]\right) = Ke^{r\delta}w_i^\delta(\theta, x)
\end{aligned}
\tag{4.44}
$$

where we used the induction hypothesis for the last inequality. It then proves that for $\theta \in [n\delta, (n+1)\delta]$ and $x \geq 0$, the right hand inequality of Equation (4.38) still holds. So we have proved by induction that it holds for $\theta, x \geq 0$.

□

Lemma 4.5.2 *For $\delta > 0$, $\theta \geq 0$, the mappings $x \mapsto x + w_i^\delta(\theta, x)$ and $x \mapsto -w_i^\delta(\theta, x)$ are non-decreasing, or equivalently, $x \mapsto w_i^\delta(\theta, x)$ is 1-Lipschitz and non-increasing. The mapping $(\theta, x) \in \mathbb{R}_+^\star \times \mathbb{R}_+ \mapsto w_i^\delta(\theta, x) \in \mathbb{R}_+$ is bounded by 1.*

Proof. For $0 \leq \theta < \delta$, since $Kw_i^\delta(\theta, x) = \max\left(K\bar{u}(\theta, x), \mathbb{E}\left[e^{-r\theta}u_i(0, K\bar{S}^x_\theta)\right]\right)$, and both $\bar{u}$ and $(\theta, x) \mapsto e^{-r\theta}\mathbb{E}\left[\frac{u_i(0, K\bar{S}^x_\theta)}{K}\right]$ are globally continuous and bounded by 1 and non-decreasing and 1-Lipschitz in the space variable, w_i^δ is continuous and bounded by 1 and non-decreasing and 1-Lipschitz in the space variable for $\theta \leq \delta$. Now, inductively, since for $\theta \geq \delta$:

$$
w_i^\delta(\theta, x) = \max\left(\bar{u}(\delta, x), e^{-r\delta}\mathbb{E}\left[w_i^\delta(\theta-\delta, \bar{S}^x_\delta)\right]\right) \tag{4.45}
$$

we get for the same reasons as previously that $x \mapsto w_i^\delta(\theta, x)$ is non-increasing and 1-Lipschitz and bounded by 1. □

Corollary 4.5.3 *For any compact set $C \subset \mathbb{R}_+$, w_i^δ tends uniformly on $\mathbb{R}_+ \times C$ to $(\theta, x) \mapsto u_i(\theta, x/K)/K$. Let $\theta, x \geq 0$:*

$$0 \leq \frac{u_i(\theta, \frac{x}{K})}{K} - w_i^\delta(\theta, x) \leq e^{r\delta} - 1 + x(1 - e^{-r\delta}) \tag{4.46}$$

Proof. Let $\theta, x \geq 0$. By Proposition 4.5.1, we get :

$$w_i^\delta(\theta, x) \leq u_i(\theta, x/K)/K \leq e^{r\delta} w_i^\delta(\theta, xe^{-r\delta}), \tag{4.47}$$

and by Lemma 4.5.2, we get :

$$\begin{aligned} 0 \leq & e^{r\delta} w_i^\delta(\theta, xe^{-r\delta}) - w_i^\delta(\theta, x) \\ &= e^{r\delta} w_i^\delta(\theta, xe^{-r\delta}) - w_i^\delta(\theta, xe^{-r\delta}) + w_i^\delta(\theta, xe^{-r\delta}) - w_i^\delta(\theta, x) \\ &\leq e^{r\delta} - 1 + x(1 - e^{-r\delta}). \end{aligned} \tag{4.48}$$

□

4.6 Appendix

Lemma 4.6.1 *Let $(\theta, x) \in \mathbb{R}_+^2$.*

The process $\left(G_t = \max(e^{-rt}(K - \bar{S}_t^x)^+, \mathbb{E}\left[e^{-r\theta} u_i(0, \bar{S}_\theta^x) | \mathcal{F}_t^X\right]\right)_{t \in [0,\theta]}$ is right-continuous in time and left-continuous over stopping times, i.e for any increasing sequence of stopping times $(\tau_n)_{n \geq 0}$ $\mathbb{P}(\lim_{n \to +\infty} G_{\tau_n} = G_{\lim_{n \to +\infty} \tau_n}) = 1$.

Proof. As $G_t = e^{-rt} K$ if $x = 0$, we assume $x > 0$. Since $v_i : x \in \mathbb{R} \mapsto u_i(0, e^x) \in \mathbb{R}_+$ is a continuous bounded function by Lemma 4.2.2, and since X is a Lévy process, by the strong Markov property $\mathbb{E}\left[e^{-r\theta} u_i(0, \bar{S}_\theta^x) | \mathcal{F}_t^X\right] = e^{-r\theta} \mathrm{P}_{\theta - t}^X v_i(\ln x + X_t)$, where $(\mathrm{P}_t^X)_{t \geq 0}$ is the transition kernel of X. By standard results on Lévy process (cf [Ber96, Prop.5.p.19]), $\tilde{v} : (t, x) \in \mathbb{R}^+ \times \mathbb{R} \mapsto \mathrm{P}_t^X v_i(x)$ is a continuous function and thus, by defining $V : (t, x) \in [0, \theta] \times \mathbb{R} \mapsto \max(e^{-rt}(K - e^x)^+, e^{-r\theta} \tilde{v}(\theta - t, x))$, one has $(G_t = V(t, \ln x + X_t))_{t \in [0,\theta]}$. Right-continuity of G is a plain consequence of the right-continuity of paths of X and left-continuity over stopping times of G is a consequence of the left-continuity over stopping times of X as a Lévy process as it is stated in [Ber96, Prop.7 p.21]. □

Part II

Dynamic programming principle with expectation constraints

5

Dynamic programming principle with conditionnal expectation constraints

Summary. We deal with stochastic optimization problems. We work in a discrete-time filtered structure, and we aim at giving general conditions about the space of controls, the cost function and the constraints to have a dynamic programming with respect to the realizations. After retrieving some well known results about the case of stochastic optimization with state constraints, by using some results of Evstigneev, we mainly provide a result which is valid with conditional expectations constraints for the control policy. To be precise, by using Balder's work, we prove under very standard assumptions about the cost function that if the filtered probability space can be reduced to the one generated by a sequence of independent random variables taking values in a Lusin space, and if the controls lie in a Polish space, and the constraints are such that at time t, the probability of the control policy must satisfy some expectation constraints conditionally to the information available at time $t-1$, then we can find the optimal strategy and the value of the optimal control by dynamic programming. This result matches exactly the purpose of general expectation constraints thanks to the works of Bouchard, Elie and Touzi.

Introduction

We address the problem of minimizing the expectation of a cost depending on the realizations of some variables with constraints about the probability laws of the controls. For short, there are controls $\mathbf{U} = (\mathbf{U}_t)_{t=0\cdots T}$ and a noise ω or $\mathbf{W} = (\mathbf{W}_t)_{t=0\cdots T}$ depending on whether or not we want to emphasize the topological properties of the underlying probability space. The cost function is canonically denoted by φ and is a function of the controls and the noise and can take value in $\mathbb{R} \cup \{+\infty\}$. We add $+\infty$ to take into account the constraints, which is very standard in optimization. There is a filtered structure on the probability space, and probability law of the noise is assumed to be known. We aim at proving that the value and the optimal control which minimize the expected cost is given by the solution of a dynamic programming equation. This problem has been extensively studied in the literature. ([Bel54, Evs76, CV77, BS78, RW98]). However, to our knowledge, no one addresses directly the problem of proving a dynamic programming principle for stochastic minimization problems with conditional expectation constraints. To illustrate what we mean by conditional expectation constraint, we give an example. Assume that at each date t the control denoted by $\mathbf{U}_t$ must take its value into $\mathbb{R}$, must be adapted to the quantity of information available at time t and has to satisfy $\mathbb{E}\left[\mathbf{U}_t^2 \middle| \mathcal{F}_{t-1}\right] \leq 1$

where $\mathcal{F}_{t-1}$ is the information available at date $t-1$. This kind of constraints is what we call conditional expectation constraints.

In §5.1, we state some results derived from those of Evstigneev [Evs76]. It gives us a unified approach to stochastic optimization problem with expectation constraints convenient to our purposes. In §5.2, we study a stochastic optimization problem with expectation constraints. By mixing the approach of the first section with results of Balder [Bal00], we prove that under fair conditions, stochastic optimization problems with conditional expectation constraints can be solved by dynamic programming. This result is very well adapted for Chapter 6.

5.1 General stochastic optimization problems under classical state constraints

The aim of this first section is to give some results about dynamic programming principle (DPP). In §5.1.1 we present the problem and the notations. In §5.1.2 we recall some results from [Evs76], which mainly provide a framework to have a measurable selection property of the optimal control. In §5.1.3, we apply the results of Evstigneev in the case where there is a state and so we find under general conditions the existence of measurable selections for controls depending on the state. These results are very well known in the stochastic optimal control community, however, as far as we know, their formal derivations from Evstigneev's results are not mentionned in the literature. Under different hypothesis, some authors have already established similar results (see [BS78, Prop.8.6]). In §5.2, we state results in the spirit of those of Evstigneev, and the formal derivation we make in §5.1.3 can thus be directly applied to the results of §5.2.

5.1.1 Notations and preliminaries

Notations

Let $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P}\right)$ be a discrete time filtered universally complete probability space such that each $\mathcal{F}_t$ is a $\mathbb{P}$-universally-complete sub σ-algebra of $\mathcal{F}$. Let $(\mathbb{U}_t)_{t\in[0,T]}$ be a collection of Polish spaces.

We recall that the universally completion of a σ-algebra $\mathcal{F}$ on Ω is the intersection of all the completions of $\mathcal{F}$ with respect to a finite measure on $(\Omega, \mathcal{F})$ for all the finite measures on $(\Omega, \mathcal{F})$. We denote by $\mathbb{U}_{0:t-1}$ the product of the spaces $\mathbb{U}_s$ from $s=0$ to $t-1$ i.e $\mathbb{U}_{0:t-1} = \prod_{i=0}^{t-1} \mathbb{U}_i$. And to simplify the notations $\mathbb{U}_{0:T-1}$ is simply denoted by $\mathbb{U}$.

The set of sequences of measurable mapping $(\mathbf{U}_t)_{t\in[0,T-1]}$ from $(\Omega, \mathcal{F})$ and taking values in $(\mathbb{U}, \mathrm{Bor}(\mathbb{U}))$ is denoted by $\mathcal{U}$.

We will use the notation $\mathbf{U}_{s:t-1}$ to denote the subsequence $(\mathbf{U}_k)_{k\in[s,t-1]}$ for $0 \le s < t < T$ and $\mathbf{U}$ will denote the whole sequence $\mathbf{U}_{0:T-1}$.

We will assume in the sequel that for a given t, the σ-algebra $\mathcal{F}_t$ is the information which has been revealed up to time t and which can be used by the optimizer at time t. In other words, the optimizer has to be non-anticipative and have to use controls given as an $\mathcal{F}_t$-adapted sequence of random variables. This lead to state optimization problems with measurability constraints. We denote by $\mathcal{U}^{\mathrm{a}}$, the subset of $\mathcal{U}$ consisting of all the $\mathcal{F}_t$-adapted sequences of randdom variable. Note that more general measurability constraints could be considered. The control at time t could be constrained to be $\mathcal{G}_t$-measurable for some σ-algebra $\mathcal{G}_t \subset \mathcal{F}_t$ without $(\mathcal{G}_t)_{t\in[0,T]}$ being a filtration.

Definition 5.1. *For $0 \le t_0 \le t_1 \le T$. Any sequence of random variables $\mathbf{Y} = (\mathbf{Y}_s)_{t_0:t_1}$ is said to be* adapted*, if*

$$\forall s \in [t_0, t_1],\ \mathbf{Y}_s \quad \textit{is } \mathcal{F}_s\textit{-measurable} .$$

We say that the sequence of random variables $\mathbf{U} \in \mathcal{U}$ is an adapted control policy *if $\mathbf{U} = (\mathbf{U}_t)_{t\in[0,T-1]}$ is such that*

$$\forall t \in [0, T-1],\ \mathbf{U}_t : (\Omega, \mathcal{F}_t) \to (\mathbb{U}_t, \mathrm{Bor}\,(\mathbb{U}_t)) \quad \textit{is } \mathcal{F}_t\textit{-measurable} .$$

The set of adapted control policies is denoted by $\mathcal{U}^{\mathrm{a}}$.

The optimization problem to be solved

In all these lines, numerical functions are valued in $\mathbb{R} \cup \{+\infty\}$, with the standard conventions. To be precise, $\mathbb{R}\cup\{+\infty\}$ is equipped with the distance $d(x, y) = |e^{-x} - e^{-y}|$ with the convention $e^{-(+\infty)} = 0$ and $x + (+\infty) = +\infty$ for any $x \in \mathbb{R} \cup \{+\infty\}$.

Let $\varphi : \mathbb{U}\times\Omega \to \mathbb{R}\cup\{+\infty\}$ be a $\mathrm{Bor}(\mathbb{U})\otimes\mathcal{F}$ measurable. The mapping φ may be thought as a cost to be minimized. The T first variables of φ correspond to the controls whereas the last one corresponds to the randomness. We are interested in minimizing the expectation of the cost by acting on it with the control sequence $\mathbf{U} \in \mathcal{U}^{\mathrm{ad}}$. The control sequence are constrained to be adapted and we therefore use as the set of admissible controls $\mathcal{U}^{\mathrm{ad}} := \mathcal{U}^{\mathrm{a}}$. As the cost may be random even when the controls are not, we will use non-standard notations. Thus, in the sequel $\mathbb{E}\left[\varphi(\mathbf{U}, \omega)\right] := \int_\Omega \varphi(\mathbf{U}(\omega), \omega)\mathbb{P}(d\omega)$. With these notations, we are interested in solving the following optimization problem:

Find $\mathbf{U}^\sharp \in \mathcal{U}^{\mathrm{ad}} = \mathcal{U}^{\mathrm{a}}$ such that:

$$\mathbb{E}\left[\varphi(\mathbf{U}^\sharp, \omega)\right] = \inf_{\mathbf{U}\in\mathcal{U}^{\mathrm{ad}}} \mathbb{E}\left[\varphi(\mathbf{U}, \omega)\right] \ . \tag{5.1}$$

Note that problems of interest are those where there exists an admissible control policy which reaches the infimum.

In the next few lines, we are going to make a quick but unrigorous reasonning to motivate the inductive definition of Equation (5.3a) and (5.3b).

Unrigorous reasoning

Let first notice that for any $\mathbf{U} \in \mathcal{U}$ such that $\mathbb{E}\left[\varphi(\mathbf{U},\omega)\right] < +\infty$ one has by properties of conditional expectations and by Fatou's lemma that:

$$\begin{aligned}\mathbb{E}\left[\varphi(\mathbf{U},\omega)\right] = & \ \mathbb{E}\left[\mathbb{E}\left[\mathbb{E}\left[\dots\mathbb{E}\left[\varphi(\mathbf{U}_0,\dots,\mathbf{U}_{T-1},\omega)|\mathcal{F}_{T-1}\right]\dots|\mathcal{F}_1\right]|\mathcal{F}_0\right]\right] \\ \geq & \ \mathbb{E}\left[\inf_{u_0\in\mathbb{U}_0}\mathbb{E}\left[\dots\inf_{u_{T-1}\in\mathbb{U}_{T-1}}\mathbb{E}\left[\varphi(u_0,\dots,u_{T-1},\omega)|\mathcal{F}_{T-1}\right]\dots\middle|\mathcal{F}_0\right]\right]\end{aligned} \tag{5.2}$$

This remark gives the idea to introduce the following problem.

$$\forall(u_0,\dots,u_{T-1})\in\mathbb{U},\quad \Phi_T(u_0,\dots,u_{T-1},\omega) = \mathbb{E}\left[\varphi(u_0,\dots,u_{T-1},\omega)|\mathcal{F}_T\right] \tag{5.3a}$$

$$\forall(u_0,\dots,u_{t-1})\in\mathbb{U}_{0:t},\ \Phi_t(u_0,\dots,u_{t-1},\omega) = \inf_{u_t\in\mathbb{U}_t}\mathbb{E}\left[\Phi_{t+1}(u_0,\dots,u_t,\omega)|\mathcal{F}_t\right]\quad \forall t\leq T-1 \tag{5.3b}$$

If we are able to ensure that there is equality in Equation (5.2), then Equation (5.3a) and (5.3b) are a way to solve Problem 5.1 under the extra condition that the policy given by these recursive constructions is well admissible.

Making Equation (5.3a) *and* (5.3b) *rigorous*

. There is already a difficulty in defining rigorously, for any φ, the mapping $u^T \in \mathbb{U} \mapsto \mathbb{E}\left[\varphi(u^T,\omega)|\mathcal{F}_{T-1}\right]$. That is why we are going to introduce the known concept of normal integrand. This concept is mentionned by many authors ([RW98, Definition 14.27 p.661], [BL73, Definition 1],[Thi81, Definition 1]). These definitions are equivalent under the assumptions that $(\Omega,\mathcal{F}_t)$ is a complete measurable space, we refer to [Roc76, Theorem 2A] and some remarks mentionned in [Roc76] and in [Vil]. Once we have noticed it, for sake of simplicity, we will use Definition 1 of [Thi81] that we recall here.

Definition 5.2. *Let* $\mathbb{F}$ *be a polish space and* $(\Omega,\mathcal{F},\mathbb{P})$ *be a* complete *probability space. For any complete* σ*-albegra* $\mathcal{G}\subset\mathcal{F}$*, we will say that a mapping* $f:\mathbb{F}\times\Omega\to\mathbb{R}\cup\{+\infty\}$ *is a* $\mathcal{G}$-lower semi-continuous (l.s.c.) integrand *or a* normal integrand *if* f *is* $\mathrm{Bor}(\mathbb{F})\otimes\mathcal{G}$*-measurable and* $\mathbb{P}$*-a.s.* $x\mapsto f(x,\omega)$ *is l.s.c.*

Other definitions of normal integrands are given in Appendix. As we will use it in the following sections, we give also the definition of Caratheodory integrand.

Definition 5.3. *Let $\mathbb{F}$ be a Polish space and $(\Omega, \mathcal{F}, \mathbb{P})$ be a* complete *probability space. For any complete σ-algebra $\mathcal{G} \subset \mathcal{F}$, we will say that a mapping $f : \mathbb{F} \times \Omega \to \mathbb{R} \cup \{+\infty\}$ is a $\mathcal{G}$-*Caratheodory integrand *if f is* $\mathrm{Bor}\,(\mathbb{F}) \otimes \mathcal{G}$*-measurable and $\mathbb{P}$-a.s. $x \mapsto f(x, \omega)$ is continuous.*

In the next Lemma (which is proved in §5.3.1), we illustrate under strong assumptions the propagation of the normality of the integrands in Equation (5.3b).

Lemma 5.1.1 *Assume that for all $t \leq T$, the Polish space $\mathbb{U}_t$ is compact and that φ is a positive $\mathcal{F}_T$-l.s.c. integrand, then for all $t \leq T$, the function Φ_t defined by Equations* (5.3a) *and* (5.3b) *is a positive $\mathcal{F}_t$-l.s.c. integrand.*

Even under these strong assumptions, it still remains to prove a measurable selection Lemma to prove that Φ_0 is a solution to Equation (5.1).

5.1.2 Recall of results of [Evs76]

In this section, we provide assumptions to solve Equation (5.1). It requires to solve intermediary problems (5.3b). The following results give sufficient conditions under which Equation (5.3a) and (5.3b) are equivalent to solve Problem 5.1.

Notation of [Evs76]

Definition 5.4. *[Evs76] Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a universally complete probability space, and $\mathcal{H}$ a $\mathbb{P}$-complete sub-σ-algebra of $\mathcal{F}$. Let $\mathbb{X}$ be a Polish space (equipped with some metric). We say that a function $f : \mathbb{X} \times \Omega \to \mathbb{R}$ belongs to $\mathbb{B}(\mathbb{X}, \mathcal{H})$ when $\mathbb{P}$-a.s.:*

(H$_1$) *f is* $\mathrm{Bor}(\mathbb{X}) \otimes \mathcal{H}$ *measurable.*

(H$_2$) *f is bounded uniformly from below with respect to the coordinates corresponding to $\mathbb{X}$ by a random variable $h : \Omega \to \mathbb{R}$ such that $\mathbb{E}\,[|h|] < +\infty$.*

(H$_3$) *For all $c \in \mathbb{R}$ and for all $\omega \in \Omega$, the set $\{x \in \mathbb{X} |\ f(x, \omega) \leq c\}$ is compact.*

Remark 5.1.2 *As compact sets are closed in metric spaces, assumption* (H$_1$) *and* (H$_3$) *implies that f is a $\mathcal{H}$-l.s.c. integrand.*

Remark 5.1.3 *Assumption* (H$_2$) *ensures us that we can choose f as a positive normal integrand without loss of generality.*

Remark 5.1.4 *If f is a $\mathcal{H}$-l.s.c. integrand and $\mathbb{X}$ is compact, then assumption* (H_3) *is satisfied since closed sets included in a compact set are compact. Assumption* (H_3) *generalizes the assumption of compactness of $\mathbb{U}$ stated in Lemma 5.1.1.*

Main result

We consider the sequences of integrands $(\varphi_t)_{t\in 0,\dots,T}$ and $(\widetilde{\varphi}_t)_{n\in 0,\dots,T-1}$ defined by backward induction as follows. We start with $\varphi_T := \varphi$ and then

$$\forall (u_{0:t}) \in \mathbb{U}_{0:t}, \quad \widetilde{\varphi}_t(u_{0:t},\omega) := \mathbb{E}\left[\varphi_{t+1}(u_{0:t},\omega)|\mathcal{F}_t\right] \quad \forall t \in \{0,\dots,T-1\} \tag{5.4a}$$

$$\forall (u_{0:t-1}) \in \mathbb{U}_{0:t-1}, \varphi_t(u_{0:t-1},\omega) := \inf_{u_t\in\mathbb{U}_t} \widetilde{\varphi}_t((u_{0:t-1},u_t),\omega) \quad \forall t \in \{0,\dots,T\} \tag{5.4b}$$

Note that in Equation (5.4b), the function φ_0 only depends on ω. In order to prove that the previous sequence of integrands is well defined, we recall some results proved in [Evs76]. We first start with some notations. Let $\mathbb{X}$,$\mathbb{Y}$ be two Polish spaces, $(\Omega,\mathcal{F})$ a measurable space and $\mathcal{H} \subset \mathcal{G}$ two sub σ-fields of the σ-field $\mathcal{F}$. For a given mapping $f : (\mathbb{X}\times\mathbb{Y})\times\Omega \to \mathbb{R}\cup\{+\infty\}$ we consider two new mappings defined as follows. The mapping $\tilde{f} : (\mathbb{X}\times\mathbb{Y})\times\Omega \to \mathbb{R}\cup\{+\infty\}$ is defined for all $(x,y) \in \mathbb{X}\times\mathbb{Y}$ and all $\omega \in \Omega$ by:

$$\tilde{f}((x,y),\omega) := \mathbb{E}\left[f((x,y),\omega)|\mathcal{H}\right] . \tag{5.5}$$

The mapping $g : \mathbb{X}\times\Omega \to \mathbb{R}\cup\{+\infty\}$ is defined for all $x \in \mathbb{X}$ and all $\omega \in \Omega$ by:

$$g(x,\omega) := \inf_{y\in\mathbb{Y}} \tilde{f}(x,y,\omega) . \tag{5.6}$$

Using Theorem 5 and Lemma 3 of [Evs76] we can now state the following theorem (Some similar result can be found in [RW98, Proposition 14.47, p.670]).

Theorem 5.1.5 *If the mapping $f : (\mathbb{X}\times\mathbb{Y})\times\Omega \to \mathbb{R}\cup\{+\infty\}$ belongs to the set $\mathbb{B}(\mathbb{X}\times\mathbb{Y},\mathcal{G})$, and there exists a* $\mathrm{Bor}(\mathbb{X})\otimes\mathcal{H}$*-measurable application $\bar{y} : \mathbb{X}\times\Omega \to \mathbb{Y}$ such that $\mathbb{P}(\forall x \in \mathbb{X} : \tilde{f}(x,\bar{y}(x,\omega),\omega) < \infty) = 1$. Then the mappings $\tilde{f}$ and g defined by Equations* (5.5) *and* (5.6) *are well defined and we have that $\tilde{f} \in \mathbb{B}(\mathbb{X}\times\mathbb{Y},\mathcal{H})$ and $g \in \mathbb{B}(\mathbb{X},\mathcal{H})$. Moreover, there exists a* $\mathrm{Bor}(\mathbb{X})\otimes\mathcal{H}$*-measurable mapping $y^\sharp : \mathbb{X}\times\Omega \to \mathbb{Y}$ such that $g(x,\omega) = \tilde{f}(x,y^\sharp(x,\omega),\omega)$.*

Proof. We will only give a sketch of proof, for a complete proof of this theorem, the reader is referred to [Evs76]. By induction, normality of the integrand is propagated through conditional expectation (cf [Thi81] and Hess, Seri) and infimum is reached because of Theorem 5.3.1 stated in Appendix and because Assumption *(H_3)* ensures compactness. Then, it is shown that there is a control satisfying the measurability constraint. $\square$

Remark 5.1.6 *If $\mathbb{P}$-a.s. $x \mapsto f(x,\omega)$ is a convex and coercive function, then f satisfies* (H_3).

A result similar to Theorem 5.1.5 is stated by T.Pennanen [PP11] under conditions (common in finance) generalizing the ones of Assumption *(H_3)*.

We will assume for all the following results that there exists an admissible control U such that $\mathbb{E}\left[\varphi(U,\omega)\right] < \infty$.

Corollary 5.1.7 *Suppose that the mapping φ is in the set $\mathbb{B}(\mathbb{U},\mathcal{F}_T)$, then the two sequences $(\varphi_t)_{t\in 0,\dots,T}$ and $(\widetilde{\varphi}_t)_{t\in 0,\dots,T-1}$ defined by equations* (5.4a) *and* (5.4b) *are well defined and we have that:*

$$\widetilde{\varphi}_t \in \mathbb{B}(\mathbb{U}_{0:t},\mathcal{F}_t) \quad \forall t \in \{0,\dots,T-1\} \tag{5.7}$$

and

$$\varphi_t \in \mathbb{B}(\mathbb{U}_{0:t-1},\mathcal{F}_t) \quad \forall t \in \{0,\dots,T\}\,. \tag{5.8}$$

Moreover, there exists a sequence $(\gamma_t^\sharp)_{t=0,\dots,T-1}$ such that for $t \in \{0,\dots,T-1\}$, $\gamma_t^\sharp : \mathbb{U}_{0:t-1}\times \Omega \to \mathbb{U}_t$ is $\mathrm{Bor}(\mathbb{U}_{0:t-1})\otimes\mathcal{F}_t$ *measurable and for which we have that:*

$$\forall u_{0:t-1} \in \mathbb{U}_{0:t-1},\ \varphi_t(u_{0:t-1},\omega) = \widetilde{\varphi}_t((u_{0:t-1},\gamma_t^\sharp(u_{0:t-1},\omega)),\omega)\,. \tag{5.9}$$

Proof. Since a finite product of Polish spaces is still a Polish space in the product topology, we apply Theorem 5.1.5 by induction, and get the result. □

Lemma 5.1.8 *Let $\mathbf{U}\in\mathcal{U}^{\mathrm{ad}}$ be a given admissible control and $(\varphi_t)_{t\in 0,\dots,T}$ be the sequence of integrands given by Equations* (5.4a) *and* (5.4b)*. Then for all $t \in \{0,\dots,T-1\}$ we have:*

$$\varphi_t(\mathbf{U}_{0:t-1},\omega) \le \mathbb{E}\left[\varphi_{t+1}(\mathbf{U}_{0:t},\omega)|\mathcal{F}_t\right]\,. \tag{5.10}$$

Moreover, the control $\mathbf{U}^\sharp$ defined by induction as follows:

$$\mathbf{U}_0^\sharp(\omega) := \gamma_0^\sharp(\omega) \quad \textit{and} \quad \mathbf{U}_{t+1}^\sharp(\omega) := \gamma_{t+1}^\sharp(\mathbf{U}_{0:t-1}^\sharp,\omega)\,, \tag{5.11}$$

is an admissible control for which previous inequality turns out to be an equality.

$$\varphi_t(\mathbf{U}_{0:t-1}^\sharp,\omega) = \mathbb{E}\left[\varphi_{t+1}(\mathbf{U}_{0:t}^\sharp,\omega)|\mathcal{F}_t\right]\,. \tag{5.12}$$

Proof. For a given $t \in \{0,\dots,T-1\}$, using the measurable mapping $\gamma_t^\sharp$, we can define an $\mathcal{F}_t$-measurable control $\mathbf{V}_t = \gamma_t^\sharp(\mathbf{U_{0:t-1}},\omega)$ and we have using the optimal property of the mapping $\gamma_t^\sharp$:

$$\varphi_t(\mathbf{U}_{0:t-1},\omega) = \widetilde{\varphi}_t((\mathbf{U}_{0:t-1},\mathbf{V}_t),\omega) \le \widetilde{\varphi}_t((\mathbf{U}_{0:t-1},\mathbf{U}_t),\omega) = \widetilde{\varphi}_t(\mathbf{U}_{0:t},\omega)\,. \tag{5.13}$$

Moreover, using the fact that the control $\mathbf{U}_{0:t}$ is $\mathcal{F}_t$-measurable, we also have by [Thi81, Proposition 13]:

$$\widetilde{\varphi}_t(\mathbf{U}_{0:t},\omega) = \mathbb{E}\left[\varphi_{t+1}(\mathbf{U}_{0:t},\omega)|\mathcal{F}_t\right] . \tag{5.14}$$

We thus have:

$$\varphi_t(\mathbf{U}_{0:t-1},\omega) \leq \mathbb{E}\left[\varphi_{t+1}(\mathbf{U}_{0:t},\omega)|\mathcal{F}_t\right] . \tag{5.15}$$

Using equation (5.11), we easily check that $\mathbf{U}^\sharp$ is an admissible control and we have:

$$\varphi_t(\mathbf{U}^\sharp_{0:t-1},\omega) = \widetilde{\varphi}_t((\mathbf{U}^\sharp_{0:t-1},\mathbf{U}^\sharp_t),\omega) = \widetilde{\varphi}_t(\mathbf{U}^\sharp_{0:t},\omega) = \mathbb{E}\left[\varphi_{t+1}(\mathbf{U}^\sharp_{0:t},\omega)|\mathcal{F}_t\right] . \tag{5.16}$$

□

Proposition 5.1.9 *Suppose that the mapping φ is in the set $\mathbb{B}(\mathbb{U},\mathcal{F}_T)$, the control $\mathbf{U}^\sharp$ defined by induction by Equation* (5.11) *is an admissible control and it is optimal for Problem* (5.1)*.*

Proof. Let $\mathbf{U} \in \mathcal{U}^{\mathrm{ad}}$ be an admissible control then using Lemma 5.1.8 we obtain that:

$$\varphi_0(\omega) \leq \mathbb{E}\left[\varphi_T(\mathbf{U}_{0:T-1},\omega)|\mathcal{F}_{T-1}\right] . \tag{5.17}$$

And thus

$$\mathbb{E}\left[\varphi_0(\omega)\right] \leq \mathbb{E}\left[\varphi_T(\mathbf{U}_{0:T-1},\omega)\right] . \tag{5.18}$$

To conclude, we use the second part of Lemma 5.1.8 to obtain that:

$$\mathbb{E}\left[\varphi_T(\mathbf{U}^\sharp_{0:T-1},\omega)\right] = \mathbb{E}\left[\varphi_0(\omega)\right] \leq \mathbb{E}\left[\varphi_T(\mathbf{U}_{0:T-1},\omega)\right] . \tag{5.19}$$

□

5.1.3 Application to classical problem with state dynamics

In this subsection, we introduce a collection of Polish spaces $(\mathbb{X}_t)_{t\in[0,T]}$ and a sequence of integrands for $t = 1,\dots,T$

$$g_t : (\mathbb{X}_{t-1}\times\mathbb{U}_{t-1})\times\Omega \to \mathbb{X}_t . \tag{5.20}$$

For a given control $\mathbf{U} \in \mathbb{U}$ and a given random variable $\mathbf{X}_0 : \Omega \to \mathbb{X}_0$, we can define a sequence of functions $(\mathbf{X}_t)_{t\in\{0,\dots,T\}}$ as follows :

$$\mathbf{X}_{t+1} : \Omega \to \mathbb{X}_{t+1} \quad \text{with} \quad \mathbf{X}_{t+1} = g_{t+1}(\mathbf{X}_t,\mathbf{U}_t,\omega) \tag{5.21}$$

Such a sequence is denoted by $\mathbf{X}$ or $\mathbf{X}^{\mathbf{U}}$ (referred as *the state controlled by* $\mathbf{U}$). The notation $\mathbf{X}^u$ is used to denote the sequence of functions generated by the constant control $u \in \mathbb{U}$. Note that these functions may not be measurable.

The filtration generated by $(\mathbf{U}, X^{\mathbf{U}})$ is denoted by

$$\mathcal{H}_t := \sigma(\mathbf{U}_{0:t-1}, X^{\mathbf{U}}_{0:t}) . \tag{5.22}$$

We make here the following assumptions:

Assumption 5.1.10 (H$_4$) *For $1 \leq t \leq T$, the integrand g_t is a $\mathcal{F}_t$-Caratheodory integrand.*

(H$_5$) *For $1 \leq t \leq T$, we have that for any $\omega \in \Omega$ and any compact set $K \subset \mathbb{X}_t$:*

$$\{(x,u) \in \mathbb{X}_{t-1} \times \mathbb{U}_{t-1} : g_t(x,u,w) \in K\} \text{ is a compact set.}$$

Lemma 5.1.11 *Let $\mathbf{X}_0$ be a given $\mathcal{F}_0$-measurable random variable and $\mathbf{U}$ a given admissible control $\mathbf{U} \in \mathcal{U}^{\mathrm{ad}}$. Under Assumptions* (H$_4$) *and* (H$_5$)*, the sequence of random variables $\mathbf{X}^{\mathbf{U}}$ generated by Equation* (5.21) *is adapted to the filtration $(\mathcal{F}_t)_{t\in\{0,\cdots,T\}}$.*

Proof. The proof is done by induction and uses standard results about Caratheodory integrands. □

For a given mapping $\psi : (\mathbb{U} \times \mathbb{X}) \times \Omega \to \mathbb{R} \cup \{+\infty\}$ we want to solve the following optimization problem:

$$\min_{\mathbf{U}\in\mathcal{U}^{\mathrm{ad}}} \mathbb{E}\left[\psi\left(\left(\mathbf{U}, \mathbf{X}^{\mathbf{U}}\right), \omega\right)\right] . \tag{5.23}$$

The next Corollary gives sufficient conditions for the existence of a solution to the minimization problem (5.23).

Corollary 5.1.12 *Let $\psi : (\mathbb{U} \times \mathbb{X}) \times \Omega \to \mathbb{R} \cup \{+\infty\}$ be a given integrand in $\mathbb{B}(\mathbb{U} \times \mathbb{X}, \mathcal{F}_T)$. Under Assumptions* (H$_4$) *and* (H$_5$)*, the integrand $\varphi : (\mathbb{U} \times \mathbb{X}) \times \Omega \to \mathbb{R} \cup \{+\infty\}$ defined by $\varphi(u,\omega) := \psi(u, \mathbf{X}^u, \omega)$ is in $\mathbb{B}(\mathbb{U}, \mathcal{F}_T)$. Moreover, there exists $\mathbf{U}^\sharp \in \mathcal{U}^{\mathrm{ad}}$ which is a solution of the minimization problem* (5.23)*.*

Proof. Since if g is l.s.c and f is continous then $g \circ f$ is l.s.c, and since closed sets included in a compact set are compact, the assumptions *(H$_4$)* and *(H$_5$)* ensure us that φ is in $\mathbb{B}(\mathbb{U}, \mathcal{F}_T)$. We can thus use Corollary 5.1.7 to conclude and obtain an optimal control $\mathbf{U}^\sharp$ defined as in Equation (5.11). □

Remark 5.1.13 *Note that* (H$_5$) *is a sufficient condition to ensure the assertion. Indeed, let us consider the following example, when $\psi : (u,x) \in \mathbb{R}^2 \mapsto \mathbf{1}_{\{x\geq 0\}} + u^2$, then ψ is not l.s.c. in (u,x), whereas for any positive mapping $g : \mathbb{R} \to \mathbb{R}^+$, $\varphi : u \mapsto \psi(u, g(u)) = 1 + u^2$ is l.s.c.*

We want, in the sequel, to specialize previous results to more specific cases and we therefore consider the new following assumptions:

(H$_6$) For $1 \le t \le T$, the integrand $g_t : (\mathbb{X}_{t-1} \times \mathbb{U}_{t-1}) \times \Omega \to \mathbb{X}_t$ is independent with respect to the σ-field $\mathcal{F}_t$.

(H$_7$) The mapping ψ does not depend on ω.

(H$_8$) The mapping ψ is given by $\psi(u, x) = \sum_{t=0}^{T-1} L_t(u_t, x_t) + K(x_T)$ with the property that for all $t < T$, $L_t \in \mathbb{B}(\mathbb{U}_t \times \mathbb{X}_t, \mathcal{F}_0)$ and $K \in \mathbb{B}(\mathbb{X}_T, \mathcal{F}_0)$.

Corollary 5.1.14 *Under the assumptions of Corollary 5.1.12 and Assumptions* (H$_6$) *and* (H$_7$)*. There exists* $\mathbf{U}^\sharp \in \mathcal{U}^{\mathrm{ad}}$ *which is a solution of the minimization problem* (5.23) *and for all* $t \in \{0, \dots, T-1\}$ *the random variable* $\mathbf{U}_t^\sharp$ *is* $\mathcal{H}_t$*-measurable (see* (5.22)*).*

Proof. We consider the three sequences of integrands $(\eta_t)_{t \in 0,\dots,T}$, $(\overline{\eta}_t)_{t \in 0,\dots,T}$, and $(\widetilde{\eta}_t)_{n \in 0,\dots,T-1}$ defined by backward induction as follows. We start with $\eta_T := \psi$ and then for all $(u_{0:t}, x_{0:t-1}) \in \mathbb{U}_{0:t} \times \mathbb{X}_{0:t-1}$ we define recursively:

$$\overline{\eta}_{t+1}(u_{0:t}, x_{0:t}, \omega) := \eta_{t+1}(u_{0:t}, (x_{0:t}, g_{t+1}(x_t, u_t, \omega))) \quad \forall t \in \{0, \dots, T-1\} \tag{5.24a}$$

$$\widetilde{\eta}_t(u_{0:t}, x_{0:t}) := \mathbb{E}\left[\overline{\eta}_{t+1}(u_{0:t}, x_{0:t}, \omega) | \mathcal{F}_t\right] . \quad \forall t \in \{0, \dots, T-1\} \tag{5.24b}$$

$$\eta_t(u_{0:t-1}, x_{0:t}) := \inf_{u_t \in \mathbb{U}_t} \widetilde{\eta}_t((u_{0:t-1}, u_t), x_{0:t}) \quad \forall t \in \{0, \dots, T\} \tag{5.24c}$$

We can also use the integrand $\varphi(u, \omega) = \psi(u, \mathbf{X}^u)$ and we consider the associated sequences $(\varphi_t)_{t \in 0,\dots,T}$ and $(\widetilde{\varphi}_t)_{n \in 0,\dots,T-1}$ defined by Equations (5.4a) and (5.4b). Let us now prove that we have the folowing equalities for all $t \in 0, \dots, T-1$:

$$\varphi_{t+1}(u_{0:t}, \omega) = \eta_{t+1}(u_{0:t}, \mathbf{X}^u_{0:t+1}) \quad \text{and} \quad \widetilde{\varphi}_t(u_{0:t}, \omega) = \widetilde{\eta}_t(u_{0:t}, \mathbf{X}^u_{0:t}) . \tag{5.25}$$

We first note that $\varphi_T(u_{0:T-1}, \omega) = \varphi(u, \omega) = \psi(u, \mathbf{X}^u) = \eta_T(u_{0:T-1}, \mathbf{X}^u_{0:T})$. Assume that we have $\varphi_{t+1}(u_{0:t}, \omega) = \eta_{t+1}(u_{0:t}, \mathbf{X}^u_{0:t+1})$, then we obtain

$$\begin{aligned} \varphi_{t+1}(u_{0:t}, \omega) &= \eta_{t+1}(u_{0:t}, (\mathbf{X}^u_{0:t}, g_{t+1}(\mathbf{X}^u_t, u_t, \omega))) \\ &= \overline{\eta}_{t+1}(u_{0:t}, \mathbf{X}^u_{0:t}, \omega) . \end{aligned}$$

Since $\mathbf{X}^u_{0:t}$ is $\mathcal{F}_t$-measurable we can in Equation (5.24a) replace $x_{0:t}$ by $\mathbf{X}^u_{0:t}$ without changing the equality. We thus have that:

$$\widetilde{\eta}_t(u_{0:t}, \mathbf{X}^u_{0:t}) := \mathbb{E}\left[\overline{\eta}_{t+1}(u_{0:t}, \mathbf{X}^u_{0:t}, \omega) | \mathcal{F}_t\right] = \mathbb{E}\left[\varphi_{t+1}(u_{0:t}, \omega) | \mathcal{F}_t\right] = \widetilde{\varphi}_t(u_{0:t}, \omega) .$$

The second part of Equality (5.25) is established. Now, minimizing both sides of Equation (5.25) with respect to u_t, we obtain the relation $\varphi_t(u_{0:t-1}, \omega) = \eta_t(u_{0:t-1}, \mathbf{X}^u_{0:t})$. Using Equation (5.24) and with the same reasoning as in Theorem 5.1.5, we obtain the existence of a mesurable sequence of mappings $\rho_t^\sharp(u_{0:t-1}, x_{0:t-1})$ such that for all $t \in \{0, \dots, T\}$:

$$\eta_t(u_{0:t-1}, x_{0:t}) = \widetilde{\eta}_t \left((u_{0:t-1}, \rho_t^\sharp(u_{0:t-1}, x_{0:t})), x_{0:t} \right) . \tag{5.26}$$

And, using the second equality in Equation (5.25), we obtain:

$$\gamma_t^\sharp(u_{0:t-1}, \omega) = \rho_t^\sharp(u_{0:t-1}, \mathcal{X}_{0:t}^u) . \tag{5.27}$$

Thus, the optimal control $\mathbf{U}^\sharp$ build by induction using Equation (5.11) can be built here with the sequence $(\rho_t^\sharp)_{t\in\{0,\dots,T-1\}}$ and we obtain that for all $t \in \{0, \dots, T-1\}$, the control $\mathbf{U}_t^\sharp$ is $\mathcal{H}_t$-measurable.
□

Corollary 5.1.15 *Under the assumptions of Corollary 5.1.14 and assumptions* (H$_8$). *There exists a control* $\mathbf{U}^\sharp \in \mathcal{U}^{\mathrm{ad}}$ *which is a solution of the minimization problem* (5.23) *and for all* $t \in \{0, \dots, T-1\}$ *the random variable* $\mathbf{U}_t^\sharp$ *is given by* $\mathbf{U}_t^\sharp = \gamma_t(X_t^\sharp)$ *where* $\mathbf{X}^\sharp = \mathbf{X}^{\mathbf{U}^\sharp}$.

Proof. We consider again the three sequences of integrands $(\eta_t)_{t\in 0,\dots,T}$, $(\overline{\eta}_t)_{t\in 0,\dots,T}$, and $(\widetilde{\eta}_t)_{n\in 0,\dots,T-1}$ defined by Equations (5.24) and we prove by induction that we have:

$$\eta_{t+1}(u_{0:t}, x_{0:t+1}) = \sum_{s=0}^{t} L_t(u_s, x_s) + V_t(x_{t+1}) . \tag{5.28}$$

Equation (5.28) is satisfied for $t = T-1$ with $V_T := K$. Assume that Equation (5.28) is satisfied for a given value of t, then we have:

$$\overline{\eta}_{t+1}(u_{0:t}, x_{0:t}, \omega) = \sum_{s=0}^{t} L_t(u_s, x_s) + V_t(g_{t+1}(x_t, u_t, \omega)) . \tag{5.29}$$

We therefore obtain:

$$\widetilde{\eta}_t(u_{0:t}, x_{0:t}, \omega) = \sum_{s=0}^{t} L_t(u_s, x_s) + \widetilde{V}_t(x_t, u_t) \quad \text{with} \quad \widetilde{V}_t(x_t, u_t) := \mathbb{E}\left[V_t(g_{t+1}(x_t, u_t, \omega)) | \mathcal{F}_t \right] . \tag{5.30}$$

And we conclude that we have:

$$\eta_t(u_{0:t-1}, x_{0:t}) = \sum_{s=0}^{t-1} L_t(u_s, x_s) + \inf_{u_t \in \mathbb{U}_t} L_t(x_t, u_t) + \widetilde{V}_t(x_t, u_t) . \tag{5.31}$$

□

5.2 General stochastic optimization problems under expectation constraints

In view of application to expectation constraints for problems with state dynamics as in §5.1.3, we make the assumption that there is a control $\mathbf{U}_T$ at time T. The space $\mathbb{U}$ stands for $\mathbb{U}_{0:T}$. As in Definition 5.1, we denote by $\mathcal{U}^{\mathrm{a}}$ the set of *adapted* control policies.

As in §5.1, we consider a cost function φ from $\mathbb{U} \times \Omega$ to $\overline{\mathbb{R}}_+$ but we assume that at each time $t = 0 \cdots T-1$ there is a d-dimensional risk-function $\gamma_t = (\gamma_{it})_{i=1\cdots d}$ with each coordinate γ_{it} from $\mathbb{U} \times \Omega$ to $\mathbb{R}$ which characterizes our risk-position from the point of view of an external regulator and which changes the set of the admissible controls in the sense of Definition 5.5. Moreover, the set of values $\mathbb{U}_t$ of the control $\mathbf{U}_t$ at each time $t = 0 \cdots T$ is a Lusin space rather than a Polish space. In the previous section, the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ had a very general structure and we have denoted $\mathbb{E}\left[\varphi(\mathbf{U}, \omega)\right] := \int_\Omega \varphi(\mathbf{U}(\omega), \omega)\mathbb{P}(d\omega)$. In this section, as we will not be able to get results with such a generality, we now introduce the canonical random variable $\mathbf{W} : \omega \mapsto \omega$ as the identity on Ω. This is to emphasize that the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ will be assumed to be not too big and not too small, in a sense that will be precised in the assumptions of Theorem 5.2.6.

Definition 5.5. *We define $\mathcal{U}^\gamma$ the set of sequence of controls satisfying the following inequalities:*

$$\forall t \in \{0, \cdots, T-1\}, \mathbb{E}\left[\gamma_t(\mathbf{U}, \mathbf{W}) | \mathcal{F}_t\right] \leq 1 \ \mathbb{P}\text{-}a.s\,.$$

In the sequel, for any $a \in \mathbb{R}^d$ and $b \in \mathbb{R}^d$, $a \leq b \Leftrightarrow b - a \in \mathbb{R}^d_+$.

As opposed to §5.1, the set of admissible control policies is now $\mathcal{U}^{\mathrm{ad}} = \mathcal{U}^{\mathrm{a}} \cap \mathcal{U}^\gamma$ and we aim at solving now the following optimization problem:

Find $\mathbf{U}^\sharp \in \mathcal{U}^{\mathrm{ad}} = \mathcal{U}^{\mathrm{a}} \cap \mathcal{U}^\gamma$ such that:

$$\mathbb{E}\left[\varphi(\mathbf{U}^\sharp, \mathbf{W})\right] = \inf_{\mathbf{U} \in \mathcal{U}^{\mathrm{ad}}} \mathbb{E}\left[\varphi(\mathbf{U}, \mathbf{W})\right] \tag{5.32}$$

We are going to work in a more specific context than in §5.1.

Assumption 5.2.1 (Ω is not too big) *The space $\Omega := \mathbb{W}_{0:T} := \mathbb{W}$ is the product of $T+1$ Lusin spaces $(\mathbb{W}_t)_{t=0\cdots T}$ and the filtration $(\mathcal{F}_t)_{t=0\cdots T}$ is the one generated by the canonical process $\mathbf{W}_t(\omega) = \omega_t$ for $t = 0 \cdots T$ and $\omega \in \Omega$, i.e $\mathcal{F}_t = \mathrm{Bor}(\mathbb{W}_{0:t})$.*

Assumption 5.2.2 (Underlying noise structure) *We assume that $\mathbb{P} = \mathbb{P}_0 \otimes \mathbb{P}_1 \otimes \cdots \otimes \mathbb{P}_T$ in the sense that for $A \in \mathrm{Bor}(\mathbb{W}_{0:T})$, $\mathbb{P}(A) = \prod_{t=0}^T \mathbb{P}_t(A_t)$. The probability $\mathbb{P}$ is the product measure of $(\mathbb{P}_t)_{t=0\cdots T}$. The probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is thus the product probability space of $(\mathbb{W}_t, \mathrm{Bor}(\mathbb{W}_t), \mathbb{P}_t)_{t=0\cdots T}$.*

Due to [Bou71, Prop.IX.68], we recall that in the case of Lusin spaces, the Borel-σ-field on the product space is the product σ-field of the Borel σ-fields. So in the sequel, for $0 \leq t \leq s \leq T$, there is no difference between the Borel σ-field of $\mathbb{W}_{t:s}$ and the product σ-field of the Borel σ-field on $\mathbb{W}_k$ for $t \leq k \leq s$, $\mathrm{Bor}(\mathbb{W}_{t:s}) = \otimes_{k=t}^{s}\mathrm{Bor}(\mathbb{W}_k)$.

Definition 5.6. *We define* $\left(\widetilde{\mathbb{W}}_{0:T}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}}\right)$ *as the augmented filtered probability space, where* $\widetilde{\mathbb{W}}_t = [0,1] \times \mathbb{W}_t$, $\widetilde{\mathbb{P}}_t = \lambda \otimes \mathbb{P}_t$ *with* λ *the Lebesgue measure on* $[0,1]$*. By analogy with the above lines, the filtration is given for any* $t = 0 \cdots T$ *by* $\widetilde{\mathcal{F}}_t = \mathrm{Bor}(\widetilde{\mathbb{W}}_{0:t})$ *and* $\widetilde{\mathcal{F}} = \widetilde{\mathcal{F}}_T$*.*

We define the set of adapted relaxed controls *as the set of adapted controls for* $\left(\widetilde{\mathbb{W}}_{0:T}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}}\right)$ *and we denote it by* $\mathrm{Rel}(\mathcal{U}^{\mathrm{a}})$*.*

We define the set of constrained relaxed controls *as the set of controls for* $\left(\widetilde{\mathbb{W}}_{0:T}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}}\right)$ *such that:*

$$\forall t \in \{0, \cdots, T-1\}, \widetilde{\mathbb{E}}\left[\gamma_t(\mathbf{U}, \mathbf{W})|\widetilde{\mathcal{F}}_t\right] \leq 1 \ \mathbb{P}\text{-}a.s\,,$$

and we denote this set by $\mathrm{Rel}(\mathcal{U}^{\gamma})$

The set $\bar{\mathbb{R}}$ is the classical notation of convex analysis.

Definition 5.7. *Let* $J : (\mathbb{U} \times \Omega, \mathrm{Bor}(\mathbb{U}) \otimes \mathcal{F}) \to (\overline{\mathbb{R}}, \mathrm{Bor}(\overline{\mathbb{R}}))$ *be a mesurable function. If for any* $\omega \in \Omega$*, and* $c \in \mathbb{R}$*, the level set* $\{u \in \mathbb{U} : J(u, \omega) \leq c\}$ *is compact (resp. closed)* J *is said to be* inf-compact with respect to $u \in \mathbb{U}$ *(resp.* J *is said to be* lower semi continuous (l.s.c for short) with respect to $u \in \mathbb{U}$*.)*

Definition 5.8. *Let* $J : (\mathbb{U} \times \Omega, \mathrm{Bor}(\mathbb{U}) \otimes \mathcal{F}) \to (\overline{\mathbb{R}}, \mathrm{Bor}(\overline{\mathbb{R}}))$ *be a l.s.c (resp. inf-compact) function with respect to* u*. If* J *is lower bounded by a function* $q : \Omega \to \overline{\mathbb{R}}$*, such that* $\mathbb{E}\left[|q(\mathbf{W})|\right] < \infty$*, then* J *is said to be a* $\mathbb{P}$*-l.s.c (resp.* $\mathbb{P}$*-inf-compact) function with respect to to* u*.*

We extend these notions to functions taking values in $\mathbb{R}^d$, by writing that $J : \mathbb{U} \times \Omega \to \bar{\mathbb{R}}^d$ is $\mathbb{P}$-l.s.c (resp. $\mathbb{P}$-inf-compact) with respect to u if each coordinate function is $\mathbb{P}$-l.s.c (resp. $\mathbb{P}$-inf-compact).

These definitions correspond exactly to the ones of what we called Evstigneev's properties and normal integrands and are used by several authors [Bal00, Thi81].

Remark 5.2.3 *We recall that due to the definition of inf-compacity, a function* $j : \mathbb{X} \to \mathbb{R} \cup \{+\infty\}$ *which is inf-compact with respect to* x *is measurable with respect to the Borel tribes on* $\mathbb{X}$ *and* $\mathbb{R} \cup \{+\infty\}$*. For a measurable function* $j : \mathbb{X} \to \mathbb{R} \cup \{+\infty\}$*, the domain of* j *is the (measurable) set of points where* j *is finite and is denoted by* $\mathrm{dom}(j)$*. When we consider a restriction of a measurable application to a smaller subset, we consider it as a measurable application for the trace tribe.*

In this context, we can state the two following results whose proofs are postponed in §5.2.3 and §5.2.4.

Proposition 5.2.4 *Assume that Assumptions 5.2.1 and 5.2.2 hold true and φ is $\mathbb{P}$-inf-compact with respect to u and for $t = 0 \cdots T-1$, γ_{it} is $\mathbb{P}$-l.s.c with respect to u. Then there exists at least one* relaxed admissible control $\mathbf{U}^\sharp \in \mathrm{Rel}(\mathcal{U}^{\mathrm{a}}) \cap \mathrm{Rel}(\mathcal{U}^\gamma)$ *which solves the minimization problem* (5.32)*.*

But Proposition 5.2.4 is not satisfying. Indeed, we are able to get a dynamic programming principle to solve the minimization problem (5.32) by keeping in mind the whole distribution of the previous controls which is the sense of the augmented dynamic programming principle in Equation (5.43). We are going to show that in some particular case, it is not necessary to keep in mind the whole distribution of the previous controls which is the sense of the augmented dynamic programming principle in Equation (5.43) to have a dynamic programming principle as in Section 5.1.

Assumption 5.2.5 *For each $t = 0 \cdots T-1$ and $i = 1 \cdots d$, the function γ_{it} only depends on $u_{0:t+1}$ and $w_{0:t+1}$, and does not depend on (u_s, w_s) for $s > t+1$.*

Theorem 5.2.6 *Assume that Assumptions 5.2.1, 5.2.2 and 5.2.5 hold true and that φ is $\mathbb{P}$-inf-compact with respect to u and the assumption that for each $t = 0 \cdots T$, the probability space $(\mathbb{W}_t, \mathrm{Bor}(\mathbb{W}_t), \mathbb{P}_t)$ is non-atomic, and the space $\mathbb{U}_t$ is a Polish space, there exists an admissible control $\mathbf{U}^\sharp \in \mathcal{U}^{\mathrm{a}} \cap \mathcal{U}^\gamma$ which solves the minimization problem* (5.32)*. This solution is obtained through the dynamic programming principle stated in Equations* (5.33)*. In particular, the value of this minimization problem is given by V_0 (see Equation* (5.33d)*).*

Let us define for $t = 0 \cdots T-1$ the following set:

$$\mathrm{Ad}(\gamma_t^{\leq 1}) := \left\{ {\scriptstyle (u_{0:t}, w_{0:t}, \mathbf{U}_{t+1}) \in \mathbb{U}_{0:t} \times \mathbb{W}_{0:t} \times L^0(\mathbb{W}_{t+1}, \mathbb{U}_{t+1})} \mid \mathbb{E}\left[\gamma_t(u_{0:t}, \mathbf{U}_{t+1}(\mathbb{W}_{t+1}), w_{0:t}, \mathbf{W}_{t+1})\right] \leq 1 \right\} , \tag{5.33a}$$

and let $\mathcal{X}_{\mathrm{Ad}(\gamma_t^{\leq 1})}$ be the characteristic function of $\mathrm{Ad}(\gamma_t^{\leq 1})$.

Let us set $V_T : \mathbb{U}_{0:T-1} \times \mathbb{W}_{0:T-1} \to \mathbb{R} \cup \{+\infty\}$ and where:

$$V_T : (u_{0:T-1}, w_{0:T-1}) = \inf_{U_T \in L^0(\mathbb{W}_T, \mathbb{U}_T)} \mathbb{E}\left[\varphi(u_{0:T-1}, \mathbf{U}_T(\mathbb{W}_T), w_{0:T-1}, \mathbf{W}_T)\right] + \mathcal{X}_{\mathrm{Ad}(\gamma_{T-1}^{\leq 1})}(u_{0:T-1}, w_{0:T-1}, \mathbf{U}_T) , \tag{5.33b}$$

and for $t = 1 \cdots T-1$,

$$V_t(u_{0:t-1}, w_{0:t-1}) = \inf_{\mathbf{U}_t \in L^0(\mathbb{W}_t, \mathbb{U}_t)} \mathbb{E}\left[V_{t+1}(u_{0:t-1}, \mathbf{U}_t(\mathbb{W}_t), w_{0:t}, \mathbf{W}_t)\right] + \mathcal{X}_{\mathrm{Ad}(\gamma_{t-1}^{\leq 1})}(u_{0:t-1}, w_{0:t-1}, \mathbf{U}_t) . \tag{5.33c}$$

and

$$V_0 = \inf_{\mathbf{U}_0 \in L^0(\mathbb{W}_0, \mathbb{U}_0)} \mathbb{E}\left[V_1\left(\mathbf{U}_0(\mathbb{W}_0), \mathbf{W}_0\right)\right] \tag{5.33d}$$

5.2.1 Results about Young measures

We now recall some results about Young measures which are the natural objects to introduce when we deal with probability kernels or Markov transition probabilities. Note that there is no need to work with universally complete probability spaces for the results of Balder [Bal00]. Let $(\Omega, \mathcal{F})$ be a measurable space and let $(\mathbb{S}, \mathrm{Bor}(\mathbb{S}))$ be a Suslin space and let $(\mathcal{M}_1(\mathbb{S}), \mathrm{Bor}(\mathcal{M}_1(\mathbb{S})))$ be the set of probability measures on $(\mathbb{S}, \mathrm{Bor}(\mathbb{S}))$ equipped with the narrow topology.

Definition 5.9 ([Bal00]). *A measurable function ν from $(\Omega, \mathcal{F})$ to $(\mathcal{M}_1(\mathbb{S}), \mathrm{Bor}(\mathcal{M}_1(\mathbb{S})))$ is called a* Young measure*. The set of Young measures is denoted by $\mathcal{R}(\Omega, \mathbb{S})$.*

Let $\mathbb{P}$ be a probability measure on $(\Omega, \mathcal{F})$.

Definition 5.10. *Let $\mathcal{G}_{\mathbb{P}}(\Omega, \mathbb{S})$ be the set of Caratheodory integrand whose absolute value is dominated uniformly in s by a $\mathbb{P}$-integrable function. For any $g \in \mathcal{G}_{\mathbb{P}}(\Omega, \mathbb{S})$ we define:*

$$I_g^{\mathbb{P}} : \nu \in \mathcal{R}(\Omega, \mathbb{S}) \mapsto \int_{\Omega} \left(\int_{\mathbb{S}} g(\omega, s)\nu(\omega)(ds) \right) \mathbb{P}(d\omega) \in \mathbb{R} \tag{5.34}$$

and we equip $\mathcal{R}(\Omega, \mathbb{S})$ with the initial topology defined by $(I_g^{\mathbb{P}})_{g \in \mathcal{G}_{\mathbb{P}}(\Omega, \mathbb{S})}$. This topology is called the $\mathbb{P}$-narrow-*topology for Young measures.*

Definition 5.11. *We say that a sequence $(\nu_n)_{n \geq 0}$ of Young measures $\mathbb{P}$-stably converges to a Young measure ν and we denote it by $\nu_n \overset{\mathbb{P}\text{-stable}}{\longrightarrow} \nu$ when:*

$$\forall g \in \mathcal{C}_b(\mathbb{S}),\ \forall A \in \mathcal{F} : \int_A \left(\int_{\mathbb{S}} g(s)\nu_n(\omega)(ds) \right) \mathbb{P}(d\omega) \to \int_A \left(\int_{\mathbb{S}} g(s)\nu(\omega)(ds) \right) \mathbb{P}(d\omega) \tag{5.35}$$

where $\mathcal{C}_b(\mathbb{S})$ is the set of bounded continuous functions from $\mathbb{S}$ to $\mathbb{R}$.

Proposition 5.2.7 (Rem.4.2 of [Bal00]) *The $\mathbb{P}$-narrow topology for Young measures is the topology of the $\mathbb{P}$-stable convergence.*

Proposition 5.2.8 (Th.6.10 of [Kal02]) *Let λ be the Lebesgue measure on $[0,1]$. Let $\hat{\mathbb{P}} := \mathbb{P} \otimes \lambda$ be the product probability measure on $(\hat{\Omega}, \hat{\mathcal{F}}) := (\Omega \times [0,1], \mathcal{F} \otimes \mathrm{Bor}([0,1]))$. And let us denote $W : (\omega, t) \in \hat{\Omega} \mapsto \omega \in \Omega$. Then to any Young measure $\nu \in \mathcal{R}(\Omega, \mathbb{S})$ we can associate a random variable U on the probability space $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}})$ such that for any $g \in \mathcal{G}_{\mathbb{P}}(\Omega, \mathbb{S})$:*

$$\hat{\mathbb{E}}\left[g(W, U)\right] = I_g^{\mathbb{P}}(\nu). \tag{5.36}$$

Proposition 5.2.9 (Purification result (Ljapunov's theorem Th.5.10 [Bal00])) *Let us assume that the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is non-atomic. Let $W : \omega \mapsto \omega$ be the identity on Ω. Let $d \in \mathbb{N}^\star$ and let $g = (g_1, \cdots, g_d) : ((\Omega \times \mathbb{S}), \mathcal{F} \otimes \mathrm{Bor}(\mathbb{S})) \to \mathbb{R}^d$ and $\nu \in \mathcal{R}(\Omega, \mathbb{S})$ such that $I_{|g|}^{\mathbb{P}}(\nu) < \infty$ then it exists a random variable $U : (\Omega, \mathcal{F}) \to (\mathbb{S}, \mathrm{Bor}(\mathbb{S}))$ such that: $I_{g_i}^{\mathbb{P}}(\nu) = \mathbb{E}\left[g_i(W, U)\right]$ for $i = 1 \cdots d$.*

Definition 5.12. *Let $h : (\Omega \times S, \mathcal{F} \otimes \mathrm{Bor}(\mathbb{S})) \to (\overline{\mathbb{R}}_+, \mathrm{Bor}(\overline{\mathbb{R}}_+))$ be a measurable function which is $\mathbb{P}$-inf-compact with respect to s. Then a subset $\mathcal{R}_0$ of $\mathcal{R}(\Omega, \mathbb{S})$ is said to be* $\mathbb{P}$-tight *if it exists such a h such that $\sup_{\nu \in \mathcal{R}_0} I_h^{\mathbb{P}}(\nu) < +\infty$.*

Theorem 5.2.10 (Th.2.2 of [Bal89]) *If $\mathcal{R}_0 \subset \mathcal{R}(\Omega, \mathbb{S})$ is $\mathbb{P}$-tight, then $\mathcal{R}_0$ is relatively compact and relatively sequentially compact for the $\mathbb{P}$-narrow topology on Young measures.*

Corollary 5.2.11 *Let $h : (\Omega \times \mathbb{S}) \to \overline{\mathbb{R}}_+$ be a $\mathbb{P}$-inf-compact function with respect to s then the level sets of I_h are compact. I.e, $I_h^{\mathbb{P}} : \nu \in \mathcal{R}(\Omega, \mathbb{S}) \to \overline{\mathbb{R}}_+$ is $\mathbb{P}$-inf-compact with respect to ν.*

Proposition 5.2.12 (Th.4.5 of [Bal00]) *Suppose that $\mathcal{F}$ is countably generated, then there exists a semimetric d on $\mathcal{R}(\Omega, \mathbb{S})$ compatible with the $\mathbb{P}$-stable convergence. The space $\mathcal{R}(\Omega, \mathbb{S})$ quotiented by the equivalence relation $x \overset{d}{=} x' \Leftrightarrow d(x, x') = 0$ is a separable metric space. We will denote it by $\mathcal{R}_{\mathbb{P}}(\Omega, \mathbb{S})$.*

Proposition 5.2.13 (Prop.2.3.3 of [CRdFV04]) *Suppose that $\mathcal{F}$ is countably generated and $\mathbb{S}$ is Polish, then the space of Young measures $\mathcal{R}_{\mathbb{P}}(\Omega, \mathbb{S})$ is a Polish space.*

5.2.2 Embedding optimization problems into a space of Young measures

Our aim in this paragraph is to describe an embedding procedure used to reformulate the optimization problem (5.32) into a problem of minimization in $\mathcal{R}(\mathbb{W}, \mathbb{U})$. We first start by considering the function $\delta : L^0(\mathbb{W}, \mathbb{U}) \to \mathcal{R}_{\mathbb{P}}(\mathbb{W}, \mathbb{U})$ defined as follows. For any random variable $\mathbf{U} \in L^0(\mathbb{W}, \mathbb{U})$, we define $\delta_{\mathbf{U}}$ as the unique element of $\mathcal{R}_{\mathbb{P}}(\mathbb{W}, \mathbb{U})$ such that for any $g \in \mathcal{G}_{\mathbb{P}}(\mathbb{W}, \mathbb{U})$, $I_g^{\mathbb{P}}(\delta_{\mathbf{U}}) = \int_{\mathbb{W}} g(\mathbf{U}(w), w) \mathbb{P}(dw)$. The range of function δ will be denoted $\mathcal{R}_{\mathbb{P}}^{Dirac}(\mathbb{W}, \mathbb{U})$ and one easily gets that

$$\mathcal{R}_{\mathbb{P}}^{Dirac}(\mathbb{W}, \mathbb{U}) \subset \mathcal{R}_{\mathbb{P}}^{\pi}(\mathbb{W}, \mathbb{U}) := \prod_{t=0}^{T} \mathcal{R}_{\mathbb{P}}(\mathbb{W}, \mathbb{U}_t) \subset \mathcal{R}_{\mathbb{P}}(\mathbb{W}, \mathbb{U}) \tag{5.37}$$

To be more precise, we embed the product space $\prod_{t=0}^{T} \mathcal{R}_{\mathbb{P}}(\mathbb{W}, \mathbb{U}_t)$ into the space of Young measures $\mathcal{R}_{\mathbb{P}}(\mathbb{W}, \mathbb{U})$ by associating to a sequence of measures $(\mu_0, \cdots \mu_T) \in \prod_{t=0}^{T} \mathcal{R}_{\mathbb{P}}(\mathbb{W}, \mathbb{U}_t)$ the product Young measure $\bigotimes_{t=0}^{T} \mu_t \in \mathcal{R}_{\mathbb{P}}(\mathbb{W}, \mathbb{U})$ defined by $\mu_0 \otimes \cdots \otimes \mu_T : w \in \mathbb{W} \mapsto \mu_0(w) \otimes \cdots \otimes \mu_T(w)$ where $\otimes$ has the usual signification of product measure. We denote $\mathcal{R}_{\mathbb{P}}^{\pi}(\mathbb{W}, \mathbb{U})$ the set of $\bigotimes_{t=0}^{T} \mu_t \in \mathcal{R}_{\mathbb{P}}(\mathbb{W}, \mathbb{U})$ for $(\mu_0, \cdots \mu_T) \in \prod_{t=0}^{T} \mathcal{R}_{\mathbb{P}}(\mathbb{W}, \mathbb{U}_t)$.

Remark 5.2.14 *Due to Proposition 5.2.7, the product topology on $\prod_{t=0}^{T} \mathcal{R}_{\mathbb{P}}(\mathbb{W}, \mathbb{U}_t)$ is coarser than the trace topology of $\mathcal{R}_{\mathbb{P}}(\mathbb{W}, \mathbb{U})$ on $\mathcal{R}_{\mathbb{P}}^{\pi}(\mathbb{W}, \mathbb{U})$. Thus, if $J : \mathcal{R}_{\mathbb{P}}(\mathbb{W}, \mathbb{U})$ is $\mathbb{P}$-inf-compact, then $J^{\pi} : (\mu_0, \cdots, \mu_T) \in \prod_{t=0}^{T} \mathcal{R}_{\mathbb{P}}(\mathbb{W}, \mathbb{U}_t) \mapsto J(\mu_0 \otimes \cdots \otimes \mu_T) \in \mathbb{R} \cup \{+\infty\}$ is $\mathbb{P}$-inf-compact for the product topology.*

Moreover, we can consider $\prod_{t=0}^{T} \mathcal{R}_{\mathbb{P}}(\mathbb{W}_{0:t}, \mathbb{U}_t)$ which is the subset of adapted sequences of Young measures and which will be denoted in the sequel by $\mathcal{R}^{\mathrm{a}}_{\mathbb{P}}(\mathbb{W}, \mathbb{U})$. This set is abusively considered as a subset of $\mathcal{R}_{\mathbb{P}}(\mathbb{W}, \mathbb{U})$ by considering that it is embedded as previously explained.

We obtain that the range of adapted sequences of random variables through function δ is the subset of adapted sequences of Young measures $\delta(\mathcal{U}^{\mathrm{a}}) \subset \mathcal{R}^{\mathrm{a}}_{\mathbb{P}}(\mathbb{W}, \mathbb{U})$. Moreover, using [Art01] we have that $\mathcal{R}^{\mathrm{a}}_{\mathbb{P}}(\mathbb{W}, \mathbb{U})$ is a closed subset of $\mathcal{R}(\mathbb{W}, \mathbb{U})$.[1]

We proceed in the same way to embed $\mathcal{U}^{\gamma}$. We define $\mathcal{R}^{\gamma}_{\mathbb{P}}(\mathbb{W}, \mathbb{U})$ as the intersection for all $t_0 = 0 \cdots T$ of the intersections of the sets

$$\bigcap_{\substack{u_{0:t_0} \in \mathbb{U}_{0:t_0} \\ w_{0:t_0} \in \mathbb{W}_{0:t_0}}} \left\{ \mu \in \mathcal{R}_{\mathbb{P}}(\mathbb{W}, \mathbb{U}) \ : \ I^{\mathbb{P}}_{\gamma_{t_0}(u_{0:t_0}, \cdot, w_{0:t_0}, \cdot)}(\mu) \leq 1 \right\} . \tag{5.38}$$

where $\gamma_{it_0}(u_{0:t_0}, \cdot, w_{0:t_0}, \cdot) : \mathbb{U}_{t_0+1:T} \times \mathbb{W}_{t_0+1:T} \to \mathbb{R} \cup \{+\infty\}$ is a convenient notation to denote the function γ_{it_0} where $(u_{0:t_0}, w_{0:t_0})$ is fixed. Note that, if for any $t = 0 \cdots T-1$, and any $i = 1 \cdots d$, the function γ_{it} is l.s.c with respect to u then these sets are closed for the narrow topology of Young measures as the intersection of closed sets. Therefore, we obtain that the set $\mathcal{R}^{\gamma}_{\mathbb{P}}(\mathbb{W}, \mathbb{U})$ is a closed subset of $\mathcal{R}(\mathbb{W}, \mathbb{U})$ when the collection of function γ are l.s.c with respect to u.

We end by embedding the cost function. Let $\varphi : \mathbb{U} \times \Omega \to \mathbb{R} \cup \{+\infty\}$ be a $\mathrm{Bor}(\mathbb{U}) \otimes \mathcal{F}$ measurable function which is lower bounded by an element h of $\mathcal{G}_{\mathbb{P}}(\mathbb{W}, \mathbb{U})$. We embed the function $J : U \in \mathcal{U}^{\mathrm{ad}} \mapsto \mathbb{E}\left[\varphi(\mathbf{U}, \omega)\right]$ into $I^{\mathbb{P}}_{\varphi} : \mathcal{R}(\mathbb{W}, \mathbb{U}) \to \mathbb{R} \cup \{+\infty\}$ by setting $I^{\mathbb{P}}_{\varphi}(\mu) = I^{\mathbb{P}}_{\varphi - h}(\mu) + I^{\mathbb{P}}_{h}(\mu)$. By definition of δ we obtain that $J(\mathbf{U}) = I^{\mathbb{P}}_{\varphi}(\delta_{\mathbf{U}})$ for all random variables $\mathbf{U} \in L^0(\mathbb{W}, \mathbb{U})$.

We can now consider the relaxed version of Problem (5.32) defined as follows:

Find $\nu^{\sharp} \in \mathcal{R}^{\mathrm{ad}}_{\mathbb{P}}(\mathbb{W}, \mathbb{U}) = \mathcal{R}^{\mathrm{a}}_{\mathbb{P}}(\mathbb{W}, \mathbb{U}) \cap \mathcal{R}^{\gamma}_{\mathbb{P}}(\mathbb{W}, \mathbb{U})$ such that:

$$I^{\mathbb{P}}_{\varphi}(\nu^{\sharp}) = \inf_{\nu \in \mathcal{R}^{\mathrm{ad}}_{\mathbb{P}}(\mathbb{W}, \mathbb{U})} I^{\mathbb{P}}_{\varphi}(\nu) \tag{5.39}$$

[1] Note that $\mathcal{R}^{\mathrm{a}}_{\mathbb{P}}(\mathbb{W}, \mathbb{U})$ is the intersection of the closed sets:

$$\left\{ \mu \in \mathcal{R}_{\mathbb{P}}(\mathbb{W}, \mathbb{U}) \ : \ I^{\mathbb{P}}_{gh}(\mu) - I^{\mathbb{P}}_{g}(\mu)\mathbb{E}\left[h(\mathbf{W}_{t+1:T})\right] = 0 \right\}$$

over $t = 0 \cdots T$, $g \in \mathcal{G}_{\mathbb{P}}(\mathbb{W}_{0:t}, \mathbb{U}_{0:t})$ and bounded Borel functions h from $\mathbb{W}_{t+1:T}$ to $\mathbb{R}$ and with the implicit convention that $\mathbb{P}$ on $\mathbb{W}_{0:t}$ is the restriction of $\mathbb{P}$ to $\mathbb{W}_{0:t}$. Moreover, one easily gets that the property for a control to be adapted depends only on the underlying filtered space and therefore does not depend on the choice of the probability measure on the whole measurable space $(\Omega, \mathcal{F})$.

5.2.3 Proof of Proposition 5.2.4 and related results

We are now able to give a proof of Proposition 5.2.4 by first proving that Problem (5.39) has a solution in $\mathcal{R}_{\mathbb{P}}(\mathbb{W},\mathbb{U})$, and therefore a relaxed solution by Proposition 5.2.8. Moreover, in some particular cases, we prove that a solution is given by a non randomized control, i.e an element in the range of δ. More precisely, we proceed as follows:

Proof. For $A \in \mathrm{Bor}(\mathcal{R}_{\mathbb{P}}(\mathbb{W},\mathbb{U}))$, we consider the characteristic function $\chi_A : \mathcal{R}_{\mathbb{P}}(\mathbb{W},\mathbb{U}) \to \mathbb{R}\cup\{+\infty\}$ of the set A defined by:

$$\chi_A(\nu) = \begin{cases} +\infty & \text{if } \nu \notin A \,, \\ 0 & \text{otherwise} \,. \end{cases} \tag{5.40}$$

and we consider the function $J_{\mathbb{P}} : \mathcal{R}_{\mathbb{P}}(\mathbb{W},\mathbb{U}) \to \overline{\mathbb{R}}$ defined by:

$$J_{\mathbb{P}}(\mu) = I^{\mathbb{P}}_{\varphi}(\mu) + \chi_{\mathcal{R}^{\mathrm{a}}_{\mathbb{P}}(\mathbb{W},\mathbb{U})}(\mu) + \chi_{\mathcal{R}^{\gamma}_{\mathbb{P}}(\mathbb{W},\mathbb{U})}(\mu) \,. \tag{5.41a}$$

where we recall that:

$$I^{\mathbb{P}}_{\varphi}(\mu) = \int_{\mathbb{W}_{0:T}} \left(\int_{\mathbb{U}_{0:T}} \varphi\left(u_{0:T}, w_{0:T}\right) \mu(w_{0:T})(du_{0:T}) \right) \mathbb{P}(dw_{0:T}) \tag{5.41b}$$

Using assumptions of Proposition 5.2.4 and the results recalled in the previous paragraph we obtain that the sets $\mathcal{R}^{\mathrm{a}}_{\mathbb{P}}(\mathbb{W},\mathbb{U})$ and $\mathcal{R}^{\gamma}_{\mathbb{P}}(\mathbb{W},\mathbb{U})$ are closed for the narrow topology of Young measures. Now, since $\mathcal{R}_{\mathbb{P}}(\mathbb{W},\mathbb{U})$ is a separable metric by Proposition 5.2.12, we obtain by Corollary 5.2.11 that $J_{\mathbb{P}}$ is $\mathbb{P}$-inf-compact.

We can use Corollary 5.3.2 to get the existence of $\mu^{\sharp} \in \mathcal{R}^{\mathrm{ad}}_{\mathbb{P}}(\mathbb{W},\mathbb{U})$ which attains the minimum of Problem (5.39) i.e $J_{\mathbb{P}}(\mu^{\sharp}) = \min_{\mu \in \mathcal{R}^{\mathrm{ad}}_{\mathbb{P}}(\mathbb{W},\mathbb{U})} J_{\mathbb{P}}(\mu)$. Using Proposition 5.2.8, we see that $\mu^{\sharp}$ is a relaxed minimizer of the minimization problem (5.32). □

We now state two results derivated from the proof of Proposition 5.2.4.

Proposition 5.2.15 *Assume that $\mathcal{R}^{\mathrm{ad}}_{\mathbb{P}}(\mathbb{W},\mathbb{U}) = \mathcal{R}^{\mathrm{a}}_{\mathbb{P}}(\mathbb{W},\mathbb{U})$ and that for any $t = 0, \cdots, T-1$, the space $\mathbb{U}_t$ is Polish, then there exists a non-randomized policy which solves the minimization problem* (5.32)*.*

Proof. We have:

$$\inf_{\mu \in \mathcal{R}^{\mathrm{a}}_{\mathbb{P}}(\mathbb{W},\mathbb{U})} J_{\mathbb{P}}(\mu) \leq \inf_{\mu \in \mathcal{R}^{\mathrm{a}}_{\mathbb{P}}(\mathbb{W},\mathbb{U}) \cap \mathcal{R}^{Dirac}_{\mathbb{P}}(\mathbb{W},\mathbb{U})} J_{\mathbb{P}}(\mu) \,.$$

Since φ is non-negative, by the same reasonning to get Equation (5.2), we have that:

$$\inf_{\mu \in \mathcal{R}^{\mathrm{a}}_{\mathbb{P}}(\mathbb{W},\mathbb{U})} J_{\mathbb{P}}(\mu) \geq \mathbb{E}\left[\inf_{u_0 \in \mathbb{U}_0} \mathbb{E}\left[\cdots \inf_{u_T \in \mathbb{U}_T} \mathbb{E}\left[\varphi(u_{0:T}, \mathbf{W}_{0:T}) \mid \mathcal{F}_T \right] \cdots \middle| \mathcal{F}_0 \right]\right] \,. \tag{5.42}$$

Thus, since by Proposition 5.1.9, we have that the right-hand side of Equation (5.42) is equal to $\inf_{\mu\in\mathcal{R}^{\mathrm{a}}_{\mathbb{P}}(\mathbb{W},\mathbb{U})\cap\mathcal{R}^{Dirac}_{\mathbb{P}}(\mathbb{W},\mathbb{U})} J_{\mathbb{P}}(\mu)$, the statement is proved. □

We now state a second result which involves the additional assumption of non-atomicity of the probability space of noises.

Proposition 5.2.16 *Assume that the probability space* $(\mathbb{W}_t, \mathrm{Bor}(\mathbb{W}_t), \mathbb{P}_t)$ *is non-atomic for* $t = 0, \cdots, T$*, then there exists a non-randomized policy which solves the minimization problem* (5.32)*.*

Before proving Proposition 5.2.16, we state a companion Lemma whose proof is postponed in §5.3.5:

Lemma 5.2.17 *Assume that the probability space* $(\mathbb{W}_T, \mathrm{Bor}(\mathbb{W}_T), \mathbb{P}_T)$ *is non-atomic. Then the function* $J_{T,\mathbb{P}}$ *defined by:*

$$J_{T-1,\mathbb{P}} : \mu_{0:T-1} \in \mathcal{R}_{\mathbb{P}}\left(\mathbb{W}, \mathbb{U}_{0:T-1}\right) \mapsto \inf_{\mu_T\in\mathcal{R}_{\mathbb{P}}(\mathbb{W},\mathbb{U}_T)} J_{\mathbb{P}}\left(\mu_{0:T-1}\otimes\mu_T\right)$$

is inf-compact with respect to $\mu_{0:T-1}$*. Moreover, there exists a non-randomized optimal control* $\mu^{\sharp}_{T}$ *which means that* $\mu^{\sharp}_{T-1}$ *is a measurable function from* $\mathrm{dom}(J_{T-1,\mathbb{P}})$ *to* $\mathcal{R}^{Dirac}_{\mathbb{P}}(\mathbb{W}, \mathbb{U}_T)$*, which satisfies that for any* $\mu_{0:T-1} \in \mathrm{dom}(J_{T-1,\mathbb{P}})$*,* $J_{T-1,\mathbb{P}}(\mu_{0:T-1}) = J_{\mathbb{P}}\left(\mu_{0:T-1}\otimes\mu^{\sharp}_{T}(\mu_{0:T-1})\right)$*.*

We prove now Proposition 5.2.16.

Proof. Using the fact that Equation (5.41a) can be written :

$$\inf_{\mu_0\in\mathcal{R}_{\mathbb{P}}(\mathbb{W},\mathbb{U}_0)}\cdots\inf_{\mu_T\in\mathcal{R}_{\mathbb{P}}(\mathbb{W},\mathbb{U}_T)} J_{\mathbb{P}}\left(\mu_0\otimes\cdots\otimes\mu_T\right)\ , \tag{5.43}$$

By a backward induction over $t = 0\cdots T$ from $t = T$ to $t = 0$, we recursively apply Lemma 5.2.17 to get that there exists a non-randomized control which solves Problem (5.32). . □

Remark 5.2.18 *Note that the general result that we have stated is that when the cost function* φ *is* $\mathbb{P}$*-inf-compact, then there exists a relaxed optimal control as soon as the admissible set is closed for the narrow topology on the Young measures. However, this does not always lead to the existence of a non randomized policy. This last point needs to be investigated carefully on each problem.*

5.2.4 One-step problem applied to partial embedding

Before proving Theorem 5.2.6, we will state Lemma 5.2.19, which is the basis element for solving Problem (5.32) in the same way as it was done for proving Theorem 5.1.5 in §5.1.

To prove Lemma 5.2.19, we will both use Theorem 5.1.5, and the concept of Young measures introduced in §5.2.1.

In this section, we use the same notation as in §5.2.2 using an horizon $T = 1$. We thus have two controls and two random variables for the noise process. We assume that the underlying spaces $\mathbb{W}$ and $\mathbb{U}$ are two product spaces of Lusin spaces. Let $\mathbb{P}$ be a Borel probability measure on the space $\mathbb{W}$ such that $\mathbb{P} = \mathbb{P}_0 \otimes \mathbb{P}_1$ where $\mathbb{P}_i$ is a probability on $\mathbb{W}_i$.

Note that in §5.2 we do not assume *contrary to assumptions in §5.1*, that the underlying probability space $(\mathbb{W}, \mathrm{Bor}(\mathbb{W}), \mathbb{P})$ is universally complete. In order to derive an extended Bellman equation which will be made more precise in §5.2.5, we want to work both with realizations $(u, w) \in \mathbb{U} \times \mathbb{W}$ and with Young measures. This will imply to deal with some new technical difficulties and some new concepts which are described now.

We will denote by $\mathrm{Univ}(\mathbb{W}_0)$ the universally completed σ-algebra of $\mathrm{Bor}(\mathbb{W}_0)$. We recall that $\mathrm{Univ}(\mathbb{W}_0)$ is the intersection of all completions of $\mathrm{Bor}(\mathbb{W}_0)$ with respect to all finite measures on $(\mathbb{W}_0, \mathrm{Bor}(\mathbb{W}_0))$. Moreover, we will work with $\mathbb{P}_0$-versions of applications. We recall that a $\mathbb{P}_0$-*version of a function* $f : \mathbb{U}_0 \times \mathbb{W}_0 \to \mathbb{R} \cup \{+\infty\}$, which is $\mathrm{Bor}(\mathbb{W}_0) \otimes \mathrm{Univ}(\mathbb{W}_0)$ measurable, is any function $\tilde{f} : \mathbb{U}_0 \times \mathbb{W}_0 \to \mathbb{R} \cup \{+\infty\}$ which is $\mathrm{Bor}(\mathbb{W}_0) \otimes \mathrm{Bor}(\mathbb{W}_0)$ measurable and such that $\mathbb{P}_0(\{w_0 \in \mathbb{W}_0 : \forall u_0 \in \mathbb{U}_0, f(u_0, w_0) = \tilde{f}(u_0, w_0)\}) = 1$.

Let $F : \mathbb{U} \times \mathbb{W} \to \mathbb{R} \cup \{+\infty\}$ be a $\mathbb{P}$-inf-compact function with respect to the pair $u \in \mathbb{U}$. And for $i = 1 \cdots d$, let γ_i be a $\mathbb{P}$-l.s.c function with respect to u.

We now embed the constraints. We set $\mathcal{X}_\gamma : (u_0, w_0, \nu_1) \in \mathbb{U}_0 \times \mathbb{W}_0 \times \mathcal{R}_{\mathbb{P}_1}(\mathbb{W}_1, \mathbb{U}_1) \to \mathbb{R}_+ \cup \{+\infty\}$ the characteristic function of the following set:

$$\left\{_{(u_0,w_0,\nu_1) \in \mathbb{U}_0 \times \mathbb{W}_0 \times \mathcal{R}_{\mathbb{P}_1}(\mathbb{W}_1,\mathbb{U}_1)} \mid \forall i = 1 \cdots d, I^{\mathbb{P}_1}_{\gamma_i(u_0,w_0,.)}(\nu_1) \le 1 \right\} . \tag{5.44}$$

and we set recall that $\mathcal{X}_{\mathcal{R}^{\mathrm{ad}}_{\mathbb{P}}(\mathbb{W},\mathbb{U})}$is the characteristic function of $\mathcal{R}^{\mathrm{ad}}_{\mathbb{P}}(\mathbb{W}, \mathbb{U})$ (see (5.39)).

The embedded problem that we consider is again:

$$\inf_{(\mu_0,\mu_1) \in \prod_{i=0}^{1} \mathcal{R}_{\mathbb{P}}(\mathbb{W},\mathbb{U}_i)} I^{\mathbb{P}}_F(\mu_0 \otimes \mu_1) + \mathcal{X}_{\mathcal{R}^{\mathrm{ad}}_{\mathbb{P}}(\mathbb{W},\mathbb{U})}(\mu_0 \otimes \mu_1) . \tag{5.45}$$

This problem can be splitted in a sequence of two optimization problems by introducing the function $\mathrm{F} : \mathcal{R}_{\mathbb{P}_0}(\mathbb{W}_0, \mathbb{U}_0) \to \mathbb{R} \cup \{+\infty\}$ as follows

$$\mathrm{F}(\mu_0) = \inf_{\mu_1 \in \mathcal{R}_{\mathbb{P}}(\mathbb{W},\mathbb{U}_1)} I^{\mathbb{P}}_F(\mu_0 \otimes \mu_1) + \mathcal{X}_\gamma(\mu_0 \otimes \mu_1) . \tag{5.46}$$

In order to get results in the spirit of §5.1, we now consider partial embedding of the function F by considering for a fixed value of the pair $(u_0, w_0) \in \mathbb{U}_0 \times \mathbb{W}_0$ the following parameterized function defined on $\mathcal{R}_{\mathbb{P}_1}(\mathbb{W}_1, \mathbb{U}_1)$:

$$I_F^{\mathbb{P}_1}(\nu_1) := \int_{\mathbb{W}_1 \times \mathbb{U}_1} F(u_{0:1}, w_{0:1}) \nu_1(w_1)(du_1) \mathbb{P}_1(dw_1) \; . \tag{5.47}$$

Note that the expression (5.47) depends parametrically on (u_0, w_0) and that the notation ν_1 in place of μ_1 is here to emphasize the fact that $\nu_1 \in \mathcal{R}_{\mathbb{P}_1}(\mathbb{W}_1, \mathbb{U}_1)$.

we denote by $f \; : \mathbb{U}_0 \times \mathbb{W}_0 \to \mathbb{R} \cup \{+\infty\}$ the parameterized function obtained by minimizing the function $I_F^{\mathbb{P}_1}$ over $\mathcal{R}_{\mathbb{P}_1}(\mathbb{W}_1, \mathbb{U}_1)$ for a fixed value of the pair $(u_0, w_0) \in \mathbb{U}_0 \times \mathbb{W}_0$:

$$f(u_0, w_0) := \inf_{\nu_1 \in \mathcal{R}_{\mathbb{P}_1}(\mathbb{W}_1, \mathbb{U}_1)} I_F^{\mathbb{P}_1}(\nu_1) + \mathcal{X}_\gamma(u_0, w_0, \nu_1) \; . \tag{5.48}$$

We will call a *solution of the parameterized minimization problem* (5.48) any application (no matter the measurability) $\nu_1^\sharp \; : \mathrm{dom}(f) \to \mathcal{R}_{\mathbb{P}_1}(\mathbb{W}_1, \mathbb{U}_1)$ such that for $(u_0, w_0) \in \mathrm{dom}(f)$, one has:

$$f(u_0, w_0) = I_F^{\mathbb{P}_1}(\nu_1^\sharp(u_0, w_0)) \; . \tag{5.49}$$

We can now embed the minimization of the function f into a minimization over $\mathcal{R}_{\mathbb{P}_0}(\mathbb{W}_0, \mathbb{U}_0)$ by considering the function $I_f^{\mathbb{P}_0}$.

In the next Lemma we prove that the above construction is valid and that we recover the solution of Problem (5.45) when minimizing f over $\mathcal{R}_{\mathbb{P}_0}(\mathbb{W}_0, \mathbb{U}_0)$.

Lemma 5.2.19 (Evstigneev's style) *Assume that the spaces $\mathbb{W}_i, \mathbb{U}_i$ are Lusin spaces for $i = 0, 1$ and assume that the space $\mathbb{U}_1$ is a Polish space. Then, f is measurable with respect to $\mathrm{Bor}(\mathbb{U}_0) \otimes \mathrm{Univ}\,(\mathbb{W}_0)$ and it exists a $\mathbb{P}_0$-version of the function f defined by Equation (5.48) which is measurable with respect to $\mathrm{Bor}(\mathbb{U}_0) \otimes \mathrm{Bor}\,(\mathbb{W}_0)$ and $\mathbb{P}_0$-inf-compact with respect to u_0. It exists a control $\nu_1^\sharp$ from $\mathbb{U}_0 \times \mathbb{W}_0$ to $\mathcal{R}_{\mathbb{P}_1}(\mathbb{W}_1, \mathbb{U}_1)$ which is $\mathrm{Bor}(\mathbb{U}_0) \otimes \mathrm{Univ}\,(\mathbb{W}_0)$ measurable and which is a solution of the minization problem (5.48). It exists also a control $\tilde{\nu}_1^\sharp$ which is $\mathrm{Bor}(\mathbb{U}_0) \otimes \mathrm{Bor}\,(\mathbb{W}_0)$ measurable and is a $\mathbb{P}_0$-version of $\nu_1^\sharp$. Moreover, we have:*

$$\inf_{\nu_0 \in \mathcal{R}_{\mathbb{P}}(\mathbb{W}_0, \mathbb{U}_0)} I_f^{\mathbb{P}_0}(\nu_0) = \inf_{\nu_0 \in \mathcal{R}_{\mathbb{P}_0}(\mathbb{W}_0, \mathbb{U}_0)} \mathrm{F}(\nu_0) = \inf_{(\nu_0, \nu_1) \in \mathcal{R}^{\mathrm{a}}_{\mathbb{P}}(\mathbb{W}, \mathbb{U})} I_F^{\mathbb{P}}(\nu_0 \otimes \nu_1) \; . \tag{5.50}$$

Before proving Lemma 5.2.19, we state this following result.

Lemma 5.2.20 *Let $\mathbb{Y}$ be a Polish space and let $f \; : \mathbb{U}_0 \times \mathbb{W}_0 \to \mathbb{Y}$ be a Borel application which is $\mathrm{Bor}(\mathbb{U}_0) \otimes \mathrm{Univ}(\mathbb{W}_0)$ measurable. Then, for any probability $\mathbb{P}_0$ on $(\mathbb{W}_0, \mathrm{Bor}(\mathbb{W}_0))$, there exists a $\mathbb{P}_0$-version of f.*

Proof. If $\mathbb{Y} = \mathbb{R} \cup \{+\infty\}$, this is exactly [Evs76, Lem.4]. The proof can be adapted to the case where the space $\mathbb{Y}$ is a Polish space as follows.

Let $d_{\mathbb{Y}}$ be the distance on the space $\mathbb{Y}$ which makes it Polish. It exists a sequence $(y_n)_{n \geq 0}$ which is dense into $\mathbb{Y}$. We set for any $\varepsilon > 0$, $N_\varepsilon \; : (u_0, w_0) \in \mathbb{U}_0 \times \mathbb{W}_0 \mapsto \inf \{n \geq 0 : d_{\mathbb{Y}}\,(f(u_0, w_0), y_n) \leq \varepsilon\}$, and define a function $f_\varepsilon\ :(u_0,w_0)\in\mathbb{U}_0\times\mathbb{Y}$ by $f_\varepsilon(u_0,w_0)=y_{N_\varepsilon(u_0,w_0)}=\sum_{n\geq 0}y_n\mathbf{1}_{\{(u_0,w_0)\in B_n\}}$. where B_0 is the preimage of the ball of center y_0 and radius $\varepsilon>0$ and for $n\geq 0$, B_{n+1} is the set of elements which lie in the preimage of the ball of center y_{n+1} and radius $\varepsilon>0$ and not in $\cup_{i=0}^n B_i$. Then, for any pair $(u_0,w_0)\in\mathbb{U}_0\times\mathbb{W}_0$ we have that $d_{\mathbb{Y}}(f_\varepsilon(u_0,w_0),f(u_0,w_0))\leq\varepsilon$. It thus proves that f_ε tends to f when ε tends to 0. Note that f_ε is a simple function in the sense of Dellacherie Meyer [DM75].

By a very general result of [Kal02, Lem.1.10] any pointwise limit of measurable functions which take values in a metric space is still measurable.

So for $\mathbb{P}_0$ on $\mathrm{Bor}(\mathbb{W}_0)$, we need only to prove that for any $n\geq 0$, and any $A\in\mathrm{Bor}(\mathbb{U}_0)\otimes\mathrm{Univ}(\mathbb{W}_0)$, there exists a $\mathbb{P}_0$-version of $y_n\mathbf{1}_{\{A\}}+y_0\mathbf{1}_{\{A^c\}}$ which is $\mathrm{Bor}(\mathbb{U}_0)\otimes\mathrm{Bor}(\mathbb{W}_0)$ measurable.

But again, by [Kal02, Lem.1.10], and the monotone class theorem, it is enough to prove that for any $B\in\mathrm{Bor}(\mathbb{U}_0)$ and $C\in\mathrm{Univ}(\mathbb{W}_0)$, it exists a $\mathbb{P}_0$-version of $y_n\mathbf{1}_{\{B\times C\}}+y_0\mathbf{1}_{\{(B\times C)^c\}}$. Since, $C\in\mathrm{Univ}(\mathbb{W}_0)$, it exists $C_1,C_2\in\mathrm{Bor}(\mathbb{W}_0)$ such that $C_1\subset C\subset C_2$ and $\mathbb{P}_0(C_1)=\mathbb{P}_0(C_2)$. Consequently, the application $y_n\mathbf{1}_{\{B\times C_2\}}+y_0\mathbf{1}_{\{(B\times C_2)^c\}}$ is $\mathrm{Bor}(\mathbb{U}_0)\otimes\mathrm{Bor}(\mathbb{W}_0)$ measurable and is equal to $y_n\mathbf{1}_{\{B\times C\}}+y_0\mathbf{1}_{\{(B\times C)^c\}}$ except for $w_0\in C_2\backslash C_1$. But $\mathbb{P}_0(C_2\backslash C_1)=0$. □

We are now able to prove Lemma 5.2.19.

Proof. Since the space $\mathbb{U}_1$ is a Polish space, we obtain using Proposition 5.2.13 that the space $\mathcal{R}_{\mathbb{P}_1}(\mathbb{W}_1,\mathbb{U}_1)$ is also Polish. Since for any $A\in\mathrm{Bor}(\mathbb{U}_0)\otimes\mathrm{Bor}(\mathbb{W}_0)$ and $B\in\mathrm{Bor}(\mathbb{W}_0)$ and $g\in\mathcal{C}_b(\mathbb{U}_1)$, the application from the space $\mathbb{U}_0\times\mathbb{W}_0\times\mathcal{R}_{\mathbb{P}_1}(\mathbb{W}_1,\mathbb{U}_1)$ and taking values in $\mathbb{R}\cup\{+\infty\}$ defined by:

$$(u_0,w_0,\nu_1)\mapsto\int_{\mathbb{W}_1}\int_{\mathbb{U}_1}\mathbf{1}_{\{A\}}(u_0,w_0)\mathbf{1}_{\{B\}}(w_1)g(u_1)\nu_1(w_1)(du_1)\mathbb{P}_1(du_1)$$

is measurable, by a density argument, one gets that $I_F^{\mathbb{P}_1}$ (See (5.47)) as a function of (u_0,w_0,ν_1) is measurable with respect to $\mathrm{Bor}(\mathbb{U}_0)\otimes\mathrm{Bor}\,(\mathbb{W}_0)\otimes\mathrm{Bor}\,(\mathcal{R}_{\mathbb{P}_1}\,(\mathbb{W}_1,\mathbb{U}_1))$. The $\mathbb{P}$-inf-compacity with respect to the pair (u_0,ν_1) follows from the fact that when $w_0\in\mathbb{W}_0$ is fixed, the application $I_F^{\mathbb{P}_1}$ from $\mathbb{U}_0\times\mathcal{R}_{\mathbb{P}_1}(\mathbb{W}_1,\mathbb{U}_1)$ is $\mathbb{P}$-inf-compact with respect to the narrow topology of Young measures on $\mathcal{R}_{\mathbb{P}}(\mathbb{W},\mathbb{U})$ which is finer than the product topology of $\mathbb{U}_0\times\mathcal{R}_{\mathbb{P}_1}(\mathbb{W}_1,\mathbb{U}_1)$. For $i=1\cdots d$, the same argument can be applied to show that the function $(u_0,w_0,\nu_1)\mapsto I^{\mathbb{P}_1}_{g_i(u_0,w_0,.)}(\nu_{t+1})$ is $\mathbb{P}$-l.s.c with respect to (u_0,ν_1) and therefore, the function $\mathcal{X}_g$ is $\mathbb{P}$-l.s.c with respect to (u_0,ν_1). Thus by Lemma 5.3.3, the function $I_F^{\mathbb{P}_1}+\mathcal{X}_g$ is $\mathbb{P}$-inf-compact with respect to (u_0,ν_1).

We can now apply Theorem 5.1.5 to the application $I_F^{\mathbb{P}_1}$ from $\mathbb{U}_0\times\mathbb{W}_0\times\mathcal{R}_{\mathbb{P}_1}\,(\mathbb{W}_1,\mathbb{U}_1)$ in order to get the measurability of function f and the existence of the measurable minimizer $\nu_1^\sharp$ (See (5.49)). At this stage, we have obtained that both functions f and $\nu_1^\sharp$ are $\mathrm{Bor}(\mathbb{U}_0)\otimes\mathrm{Univ}(\mathbb{W}_0)$ measurable.

Since both functions f and $\nu_1^\sharp$ take values in Polish spaces, the existence of $\mathbb{P}_0$-versions of f and $\nu_1^\sharp$ is ensured by Lemma 5.2.20. We denote these $\mathbb{P}_0$-versions by $\tilde{f}$ and $\tilde{\nu}_1^\sharp$.

Then, by applying the same reasonning to the function $\tilde{f}$, we get the existence of a $\mathrm{Bor}(\mathbb{W}_0)$ measurable application $\nu_0^\sharp$ from $\mathbb{W}_0$ to $\mathbb{U}_0$ such that $\tilde{f}(\nu_0^\sharp(w_0), w_0)$ is a $\mathbb{P}_0$-version of $\inf_{u_0 \in \mathbb{U}_0} \tilde{f}(u_0, w_0)$. It is then easy to check by Fatou's lemma that the infimum with function $\tilde{f}$ is not bigger than the infimum with function F.

Since the function $\delta_{\nu_0^\sharp} \in \mathcal{R}_{\mathbb{P}}^{Dirac}(\mathbb{W}_0, \mathbb{U}_0)$, the infimum with F is not greater than the infimum with $\tilde{f}$. Consequently, these two infima are equal. □

5.2.5 Application to multi-stage problems and proof of Theorem 5.2.6

We are now able to prove Theorem 5.2.6, which states that there exists an admissible control $\mathbf{U}^\sharp$ which solves the minimization problem (5.32). This solution is obtained through the dynamic programming principle stated in Equations (5.33). In Equations (5.51), we write Equations (5.33) with the formalism of Young measures.

Proof.

Let us define the following functions and sets.

$$V_{T+1} : (u_{0:T}, w_{0:T}) \in \mathbb{U} \times \mathbb{W} \mapsto \varphi(u_{0:T}, w_{0:T}) \tag{5.51a}$$

For $t = 0 \cdots T-1$, we set $\mathcal{X}_{\gamma_t}$ the characteristic function of the following set:

$$\left\{ {}_{(u_{0:t}, w_{0:t}, \nu_{t+1}) \in \mathbb{U}_{0:t} \times \mathbb{W}_{0:t} \times \mathcal{R}_{\mathbb{P}_{t+1}}(\mathbb{W}_{t+1}, \mathbb{U}_{t+1})} \mid \forall i = 1 \cdots d, I^{\mathbb{P}_{t+1}}_{\gamma_{it}(u_{0:t}, w_{0:t}, .)}(\nu_{t+1}) \leq 1 \right\} , \tag{5.51b}$$

and for $O = 1 \cdots T$, where by convention $\mathcal{X}_{\gamma_{-1}} \equiv 0$, we set:

$$V_t(u_{0:t-1}, w_{0:t-1}) = \inf_{\nu_t \in \mathcal{R}_{\mathbb{P}_t}(\mathbb{W}_t, \mathbb{U}_t)} I^{\mathbb{P}_t}_{V_{t+1}(u_{0:t-1}, w_{0:t-1}, .)}(\nu_t) + \mathcal{X}_{\gamma_{t-1}}(u_{0:t-1}, w_{0:t-1}, \nu_t) . \tag{5.51c}$$

For each $t = 0 \cdots T$, we are going to prove that there exists a $\mathbb{P}$-version of V_t which is $\mathrm{Bor}(\mathbb{W}_{0:t}) \otimes \mathrm{Bor}(\mathbb{U}_{0:t})$ measurable and $\mathbb{P}_{0:t}$-inf-compact with respect to $u_{0:t}$. This statement is true for V_{T+1}. Let $t = 0 \cdots T$, by the induction hypothesis on V_{t+1} and using Equation (5.51c), we can apply Lemma 5.2.19 with V_{t+1} in place of F, γ_{t-1} in place of γ, $\mathbb{U}_{0:t-1}, \mathbb{W}_{0:t-1}, \mathbb{P}_{0:t-1}$ in place of $\mathbb{U}_0, \mathbb{W}_0, \mathbb{P}_0$, $\mathbb{U}_t, \mathbb{W}_t, \mathbb{P}_t$ in place of $\mathbb{U}_1, \mathbb{W}_1, \mathbb{P}_1$. We then get the $\mathbb{P}_{0:t}$-version of V_t which is $\mathrm{Bor}(\mathbb{W}_{0:t}) \otimes \mathrm{Bor}(\mathbb{U}_{0:t})$ measurable and $\mathbb{P}_{0:t}$-inf-compact with respect to $u_{0:t}$, and the existence of a $\mathbb{P}_{0:t}$-version of the mesurable control $\nu_t^\sharp : \mathrm{dom}(V_t) \subset \mathbb{U}_{0:t-1} \times \mathbb{W}_{0:t-1} \to \mathcal{R}_{\mathbb{P}}(\mathbb{W}_t, \mathbb{U}_t)$ which solves the minimization problem (5.51c). If we assume moreover that $(\mathbb{W}_t, \mathrm{Bor}(\mathbb{W}_t), \mathbb{P}_t)$ is a non-atomic probability space, then by Proposition 5.2.9, and by Lemma 5.2.20, there exists a non-randomized optimal control $\mathbf{U}_t^\sharp$ which solves minimization problem (5.51). □

Note that we can solve the relaxed problem (5.39) by the mean of the dynamic programming (5.51).

5.2.6 Application to optimization problems with expectation constraints and state dynamics

In this subsection, we use the same model as in the §5.1.3. For a given mapping $\psi : \mathbb{X} \times \mathbb{U} \times \mathbb{W} \to \mathbb{R} \cup \{+\infty\}$, we want to solve the following optimization problem which is to find a control $\mathcal{U}^\sharp \in \mathcal{U}^{\mathrm{ad}} = \mathcal{U}^{\mathrm{a}} \cap \mathcal{U}^{\gamma}$ (see Definition 5.5) such that :

$$\mathbb{E}\left[\psi\left(\mathbf{U}^\sharp, \mathbf{X}^{\mathbf{U}^\sharp}, \mathbf{W}\right)\right] = \min_{\mathbf{U}\in\mathcal{U}^{\mathrm{ad}}} \mathbb{E}\left[\psi(\mathbf{U}, \mathbf{X}^{\mathbf{U}}, \mathbf{W})\right] \ . \tag{5.52}$$

Corollary 5.2.21 *Let ψ be a $\mathbb{P}$-inf-compact function with respect to (u, x). Under Assumption 5.1.10, and the assumptions of Theorem 5.2.6, there exists an admissible solution $\mathbf{U}^\sharp$ to the minimization problem* (5.52)*.*

Proof. By using the first statement of Corollary 5.1.12, we get that the function $\varphi : (u, w) \in \mathbb{U}\times\mathbb{W} \mapsto \mathbb{E}\left[\psi(u, \mathbf{X}^u, \mathbf{W}) \mid \mathbf{W} = w\right]$ is $\mathbb{P}$-inf-compact with respect to u. We now apply Theorem 5.2.6 to get the result. □

5.3 Appendix

5.3.1 Auxiliary statements of §5.1.1

Other definitions of normal integrand

Definition 5.13. *[BL73, Definition 1] Assume that the spaces Ω and $\mathbb{U}$ are locally compact Polish spaces.*
A $\mathrm{Bor}(\mathbb{U}) \times \mathcal{F}$ measurable borelian function $\varphi : \mathbb{U} \times \Omega \to \mathbb{R}$ is a ***Caratheodory function*** *if $\mathbb{P}$-a.s. it takes value in $\mathbb{R}_+$ and $(u_0, \ldots, u_{t-1}) \mapsto \varphi(u_0, \ldots, u_{t-1}, \omega)$ is continuous.*
A $\mathrm{Bor}(\mathbb{U}) \times \mathcal{F}$ measurable borelian function $\varphi : \mathbb{U} \times \Omega \to \mathbb{R}$ is a ***positive normal integrand*** *if it exists a sequence $(\varphi_n)_{n\in\mathbb{N}}$ of Caratheodory functions such that $\mathbb{P}$-a.s. $\varphi = \sup_{n\in\mathbb{N}} \varphi_n$.*

Definition 5.14. *[RW98, Definition 14.27 p.661] Let $(\mathrm{T}, \mathcal{T})$ be a measurable space. A function $f : \mathrm{T} \times \mathbb{R}^n \to \bar{\mathbb{R}}$ will be a* ***normal integrand*** *if its epigraphical mapping $S_f : \mathrm{T} \to \mathbb{R}^n \times \mathbb{R}$ is closed-valued and measurable.*

Proof of Lemma 5.1.1

Proof. We prove the lemma by induction. Using assumptionsn we obtain that $\mathbb{P}$-a.s. the function $\varphi = \Phi_T$ is lower semi continuous with respect to $\mathbb{U}$.

Now, assume that $\mathbb{P}$-a.s. the function Φ_{t+1} is l.s.c with respect to $\mathbb{U}_{t+1}$, then using [Thi81, Proposition 12], it exists $\bar{\Omega}$, such that $\mathbb{P}(\bar{\Omega}) = 1$ and such that the function $\forall \omega \in \bar{\Omega}$, $(u_0, \ldots, u_t) \mapsto \mathbb{E}\left[\Phi_{t+1}(u_0, \ldots, u_t, \omega)|\mathcal{F}_t\right]$ is l.s.c.. It remains to prove that the function $\forall \omega \in \bar{\Omega}$, $(u_0, \ldots, u_{t-1}) \mapsto \inf_{u_t \in \mathbb{U}} \mathbb{E}\left[\Phi_{t+1}(u_0, \ldots, u_t, \omega) \,|\, \mathcal{F}_t\right]$ is also l.s.c. This is not an easy result as it is illustrated in [RW98, Proposition 14.47,p. 670] (with (t, x, u) replacing $(\omega, (u_0, \ldots, u_{t-1}), u_t)$).[2] Let $\omega \in \bar{\Omega}$, and $(u_0, \ldots, u_{t-1}) \in \mathbb{U}_t$ and $(U_n^t)_{n \in \mathbb{N}}$ be a sequence converging to $(u_0, \ldots, u_{t-1})$. Let $\varepsilon > 0$, then $\forall n \geq 0$, it exists $u_n^\varepsilon \in \mathbb{U}$, such that $\Phi_t(U_n^t, \omega) \geq \mathbb{E}\left[\Phi_{t+1}(U_n^t, u_n^\varepsilon, \omega)|\mathcal{F}_t\right] - \varepsilon$. As $\mathbb{U}_t$ is compact, the set of cluster points of u_n^ε denoted by S^ε is not empty. Using l.s.c. property of function $\mathbb{E}\left[\Phi_{t+1}(\bullet, \ldots, \bullet, u, \omega)|\mathcal{F}_t\right]$ we obtain

$$\liminf_n \Phi_t(U_n^t, \omega) \geq \inf_{u \in S^\varepsilon} \mathbb{E}\left[\Phi_{t+1}(u_0, \ldots, u_{t-1}, u, \omega)|\mathcal{F}_t\right] - \varepsilon \geq \Phi_t(u_0, \ldots, u_{t-1}) - \varepsilon \ .$$

Which ends the proof. □

Properties of l.s.c functions

Definition 5.15. *Let $(\mathbb{X}, d)$ be a metric space. A function $f : \mathbb{X} \to \mathbb{R}$ is* lower semi continuous *(l.s.c) if the set $f^{\leq c} := \{x \in \mathbb{X} \mid f(x) \leq c\}$ is closed for all $c \in \mathbb{R}$.*

The following Theorem is essential whose proof can be found in classical textbooks as those of Jean Charles Gilbert, or Guy Cohen.

Theorem 5.3.1 *Let $\mathbb{X}$ be a separated topological space and let K be a compact subset of $\mathbb{X}$. Let $f : \mathbb{X} \to \mathbb{R} \cup \{+\infty\}$ be a lower semi continuous function then the function f is lower bounded on K and attains its minimum in K.*

Proof. We first recall, that if K is a compact subset in a separated topological space then for any family of closed sets $(F_i)_{i \in I}$ of $\mathbb{X}$ such that any finite intersection of $F_i \cap K$ is non empty then $\cap_{i \in I} F_i \cap K$ is non empty.

We now prove by contradiction that the function f is lower bounded. Assume that the function f is not lower bounded and consider the family of closed sets $(f_K^{\leq r})_{r \in \mathbb{R}}$ defined as follows:

[2] Example: let f be a continuous non-negative function, let us define $g : x \mapsto \inf_{y \in \mathbb{R}} \exp\{-y f(x)\} = \mathbf{1}_{\{f(x)=0\}}$ and the indicator function of a closed subset is u.s.c.

$$f_K^{\leq r} := \{x \in \mathbb{X} \mid f(x) \leq r\} \cap K \quad \forall r \in \mathbb{R} \,.$$

Then, for any $n \in \mathbb{N}$ and any sequence of real numbers $r_1, \ldots, r_n \in \mathbb{R}$, we have that $\cap_{i=1}^n f_K^{\leq r_i} = f_K^{\leq \inf_{i=1,\cdots,n} r_i} \neq \emptyset$. Consequently, there exists $x \in K$ such that $f(x) = -\infty$, which contradicts the fact that the function f takes value in $\mathbb{R} \cup \{+\infty\}$.

We can now assume that f is lower bounded on K and we can therefore define $c = \inf_{x \in K} f(x)$. For any real number $\varepsilon > 0$, the set $f_K^{\leq c+\varepsilon}$ is non empty and thus for any $n \in \mathbb{N}$ and $\varepsilon_1, \ldots, \varepsilon_n > 0$, $\cap_{i=1}^n f_K^{\leq c+\varepsilon_i} = f_K^{\leq c+\min_{i=1,\cdots,n} \varepsilon_i}$ is also non empty. Thus, we have that the set $\cap_{\varepsilon>0} f_K^{\leq c+\varepsilon}$ is non empty which proves that there exists $x^\sharp \in K$ such that $f(x^\sharp) = \inf_{x \in K} f(x)$. □

Corollary 5.3.2 *Let $\mathbb{X}$ be a separated topological space and let $f : \mathbb{X} \to \mathbb{R} \cup \{+\infty\}$ be a proper*[3] *function such that for any real number c, the set $f^{\leq c}$ is a compact subset of $\mathbb{X}$. Then, the function f attains its minimum in $\mathbb{X}$.*

Proof. Since f is proper there exists a real number c which is such that $f^{\leq c}$ is non empty. The proof follows the from Theorem 5.3.1 with $K = f^{\leq c}$. □

5.3.2 Basic lemmas

Lemma 5.3.3 *Let $\mathbb{X}$ be a metric space and let $f : \mathbb{X} \to \mathbb{R} \cup \{+\infty\}$ be an inf-compact function with respect to x and $g : \mathbb{X} \to \mathbb{R}_+ \cup \{+\infty\}$ be a l.s.c. function with respect to x. Then the sum function $f + g$ is an inf-compact function with respect to x.*

Proof. Let $c \in \mathbb{R}$, due to the non-negativity of g, one has:

$$(f+g)^{\leq c} := \{x \in \mathbb{X} : f(x) + g(x) \leq c\} \subset \{x \in \mathbb{X} \ : \ f(x) \leq c\} = f^{\leq c}$$

As $(f+g)^{\leq c}$ is a closed set because $f+g$ is a l.s.c function as the sum of two l.s.c functions and since $f^{\leq c}$ is a compact set, we have that $(f+g)^{\leq c}$ is a compact set too. □

5.3.3 A measurable selection theorem

Definition 5.16 ([CV77]). *Let $(\Omega, \mathcal{F})$ be a measurable space and X a separable metric space. The set of the sets of X denoted by 2^X is equipped with the topology of the Hausdorff's distance and 2^X will be then viewed as a measurable space with the corresponding Borel σ-algebra. An application*

[3] We say that a function is proper when it takes value on $\mathbb{R} \cup \{+\infty\}$ and is not indentically equal to $+\infty$.

$F : \Omega \to 2^X$ *is called a* multi-application *and can be also denoted by* $F : \Omega \rightrightarrows X$. *When the application* F *is measurable with respect to* $\mathcal{F}$ *and the Borel* σ*-algebra of* 2^X, F *is called a* measurable multi-application.

Theorem 5.3.4 (Kuratowski-Ryll Nardzewski's theorem [KRN65]) *Let* $(\Omega, \mathcal{F})$ *be a measurable space and* X *a separable metric space. Let* $F : \Omega \to 2^X$ *be a multi-application with complete non-empty values satisfying that for any closed set* $V \subset X$*:*

$$\{\omega \in \Omega : F(\omega) \cap V \neq \varnothing\} \in \mathcal{F} \tag{5.53}$$

then it exists a measurable selector $f : (\Omega, \mathcal{F}) \to (X, \mathrm{Bor}(X))$ *such that for any* $\omega \in \Omega$, $f(\omega) \in F(\omega)$.

The property (5.53) stated *for any open set* V rather than *for any closed set* V is the definition of the measurability of the multi-application F.

5.3.4 A measurability selection theorem for partial minimization

In the proof of Proposition 5.2.4 we have solved the problem by minimizing the function $J_{\mathbb{P}}$ in the space of Young measures $\mathcal{R}_{\mathbb{P}}(\mathbb{W}, \mathbb{U})$. It is possible to obtain the same existence result by solving a family of parameterized optimization problems in the spirit of what was done in §5.1. As opposed to §5.1 the problems are parameterized by Young measures and not by elements of the control space $\mathbb{U}$. The assumptions are also slightly different. The underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$ needs not to be complete since measurability will be deduced from l.s.c properties and we will work with separable metric spaces rather than Polish spaces.

We now state a one step problem to be solved in order to solve the global problem.

Proposition 5.3.5 *Let* $\mathbb{X}$ *and* $\mathbb{U}$ *be two separable metric spaces and assume that the measurable function* $J : \mathbb{X} \times \mathbb{U} \to \mathbb{R} \cup \{+\infty\}$ *is inf-compact with respect to the pair* (x, u) *and let us define a function* $j : \mathbb{X} \to \mathbb{R} \cup \{+\infty\}$ *as follows:*

$$j : x \in \mathbb{X} \mapsto \inf_{u \in \mathbb{U}} J(x, u) \; . \tag{5.54}$$

Then, the function j *is inf-compact with respect to* x *and it exists a measurable function* $U^\sharp : \mathrm{dom}(j) \to \mathbb{U}$ *such that* $j(x) = J(x, U^\sharp(x))$ *for all* $x \in \mathrm{dom}(j)$.

Remark 5.3.6 *The existence of the minimizer* $U^\sharp$ *in the next proposition 5.3.5 is in fact a direct consequence of [Rie78, Th.4.1]. However, the property that* j *is inf-compact with respect to* x *is not stated in [Rie78], and thus Proposition 5.3.5 is proved in the next lines.*

Proof. For $x \in \mathbb{X}$ such that $x \in \operatorname{dom} j$ the function $J(x, \cdot)$ is proper and the function $J(x, \cdot)$ is inf-compact. Thus, we can use Corollary 5.3.2 to obtain that the minimum of function $J(x, \cdot)$ is attained.

We now prove that the function j is inf-compact. Let $c \in \mathbb{R}$, and consider the set $j^{\leq c} := \{x \in \mathbb{X} \,|\, j(x) \leq c\}$. Assume that $j^{\leq c}$ is not empty and consider $(x_n)_{n\geq 0}$ a sequence of elements of $j^{\leq c}$. Then, there exists a sequence $(u_n)_{n\geq 0}$ of elements of $\mathbb{U}$ which are such that $J(x_n, u_n) \leq c$. Indeed, for all $n \geq 0$ $x_n \in \operatorname{dom} j$ and we can consider the value u_n which attains the minimum of $J(x_n, \cdot)$. The sequence $(x_n, u_n)_{n\geq 0} \in J^{\leq c}$ is a sequence in a compact set and consequently there exists $(x, u) \in J^{\leq c}$ and a subsequence $(x_{n_k}, u_{n_k})_{k\geq 0} \in J^{\leq c}$ converging to (x, u). Since for all $u \in \mathbb{U}$ we have $j(x) \leq J(x, u)$ we obtain that x and the subsequence $(x_{n_k})_{k\geq 0}$ both belongs to $j^{\leq c}$. We have found a subsequence of the sequence $(x_n)_{n\geq 0}$ converging to $x \in j^{\leq c}$ and conclude that $f^{\leq c}$ is a compact set in the metric space $\mathbb{X}$.

Now, it is clear that the multi-application $G : x \in \operatorname{dom}(j) \mapsto \{u \in \mathbb{U} : J(x, u) \leq j(x)\}$ is a multi-application with complete non-empty values since $G(x)$ is non-empty and complete as a compact set in a metric space. Now, let F be a closed set of $\mathbb{U}$, let us set $J_F(x, u) = J(x, u)$ if $u \in F$ and $+\infty$ otherwise, and let us define $j_F : x \in \operatorname{dom}(j) \mapsto \inf_{u\in\mathbb{U}} J_F(x, u)$. Since J_F keeps on being inf-compact with respect to the pair, then by the previous reasoning, j_F is inf-compact and in particular measurable. Therefore we have:

$$\{x \in \operatorname{dom}(j) : G(x) \cap F \neq \varnothing\} = \{x \in \operatorname{dom}(j) : j_F(x) = j(x)\} \tag{5.55}$$

and this set is measurable. We apply the Kuratowski-Ryll-Nardzewski's theorem (Theorem 5.3.4) to get the existence of $U^\sharp$ as stated. □

5.3.5 Proof of Lemma 5.2.17

In these few lines, we are going to prove Lemma 5.2.17. For an easier reading of the proof, we will state our problem with only two time steps, i.e $T = 2$. We proceed as in §5.2.2.

Lemma 5.3.7 *We now define:*

$$J_{0,\mathbb{P}} : \mu_0 \in \mathcal{R}_{\mathbb{P}}(\mathbb{W}, \mathbb{U}_0) \mapsto \inf_{\mu_1 \in \mathcal{R}_{\mathbb{P}}(\mathbb{W}, \mathbb{U}_1) : \mu_0 \otimes \mu_1 \in \mathcal{R}^{\mathrm{a}}_{\mathbb{P}}(\mathbb{W}, \mathbb{U})} J_{\mathbb{P}}(\mu_0 \otimes \mu_1) \tag{5.56}$$

$J_{0,\mathbb{P}}$ is inf-compact with respect to μ_0 and admits a measurable solution $\mu_1^\sharp$ from $\operatorname{dom}(J_{0,\mathbb{P}}) \subset \mathcal{R}_{\mathbb{P}}(\mathbb{W}, \mathbb{U}_0)$ to $\mathcal{R}_{\mathbb{P}}(\mathbb{W}, \mathbb{U}_1)$ such that for $\mu_0 \in \operatorname{dom}(J_{0,\mathbb{P}})$, one has $J_{0,\mathbb{P}}(\mu_0) = J_{\mathbb{P}}(\mu_0 \otimes \mu_1^\sharp(\mu_0))$.

Proof. By Remark 5.2.14, the application $J_{\mathbb{P}} : (\mu_1, \mu_2) \mapsto J_{\mathbb{P}}(\mu_1 \otimes \mu_2)$ is inf-compact for the product topology on $\mathcal{R}_{\mathbb{P}}(\mathbb{W}, \mathbb{U}_0) \times \mathcal{R}_{\mathbb{P}}(\mathbb{W}, \mathbb{U}_1)$. Then by Lemma 5.3.3, the application $J_{\mathbb{P}} + \mathcal{X}_{\mathcal{R}^{\mathrm{ad}}_{\mathbb{P}}(\mathbb{W},\mathbb{U})}$ is inf-compact. Endly, we apply Proposition 5.3.5 to the function $J_{\mathbb{P}} + \mathcal{X}_{\mathcal{R}^{\mathrm{ad}}_{\mathbb{P}}(\mathbb{W},\mathbb{U})}$. □

We are now able to prove Lemma 5.2.17. **Proof.** Since for any $\mu_0 \in \mathrm{dom}(J_{0,\mathbb{P}})$, one has the existence of $\mu_1^\sharp(\mu_0) \in \mathcal{R}_{\mathbb{P}}^{\mathrm{ad}}(\mathbb{W},\mathbb{U})$, and since by Proposition 5.2.9, one can find an element $\mathbf{U}_1^\sharp(\mu_0)$ of $\mathcal{R}_{\mathbb{P}}^{Dirac}(\mathbb{W},\mathbb{U}_1)$ such that $I_g^{\mathbb{P}}(\mathbf{U}_1^\sharp(\mu_0)) = I_g^{\mathbb{P}}(\mu_1^\sharp(\mu_0))$, one can now restate the minimization problem (5.56) with adding the closed constraint that $\mu_1 \in \mathcal{R}^{Dirac}(\mathbb{W},\mathbb{U}_1)$. So we proved that there exists a measurable application $\mathbf{U}_1^\sharp : \mu_0 \in \mathrm{dom}(J_{0,\mathbb{P}}) \mapsto \mathbf{U}_1(\mu_0) \in \mathcal{R}_{\mathbb{P}}^{Dirac}(\mathbb{W},\mathbb{U}_1)$ such that $J_{0,\mathbb{P}}(\mu_0) = J_{\mathbb{P}}(\mu_0 \otimes \mathbf{U}_1^\sharp(\mu_0))$. □

Doob's Lemma

We refer the interested reader to [Pra90] and the remark below [Vil, Th.II.32].

Lemma 5.3.8 *Let* $(\mathbb{X},\mathcal{F})$ *be a measurable space and* let $\mathbb{F}$ be a Polish space. *Let* $\mathbf{X}$ *be a* $(\mathbb{X},\mathcal{F})$*-valued random variable. Then,* $\mathbf{Y}$ *is a* $\mathbb{F}$*-valued* $\sigma(\mathbf{X})$*-measurable random variable if and only if there exists a measurable function* $h : (\mathbb{X},\mathcal{F}) \to (\mathbb{F},\mathrm{Bor}(\mathbb{F}))$ *such that* $\mathbf{Y} = h(\mathbf{X})$.

6

Optimization problems under expectation constraints

Summary. In this Chapter, we assume that we are in the framework of §5.1.3. We make the additionnal hypothesis that the underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$ has the same structure as in §5.2. Through the ideas of Bouchard, Elie, and Touzi Soner on stochastic target problems with controlled loss, we show that a problem with expectations constraints can be reduced to a problem with constraints of conditional expectations. We thus prove a dual version of the extended dynamic programming equation for standard expectation constrained problems like the case of management of the production of an hydro-electric power plant with a probability constraint on the level of the dam on certain dates.

Introduction

A practical problem of management of a dam with a constraint on the probability of tank level at a future date can be formulated as a minimization problem below (5.52). This problem is part of the class of optimal control problems in discrete time stochastic constraints with expectations. This has been studied in [CCC$^+$11] and in [Gir10, Chap.5]. These authors were particularly interested in the dynamic consistency problems of this type and their numerical resolution. The link between stochastic optimization problems with expectation constraints and those with conditional expectation constraints was made by Bouchard, Elie and Touzi [BET10], whose work was for the case of a stochastic target problem with a controlled loss in a continuous time diffusion framework.

In §6.1, thanks to the results of Theorem 5.2.6, we prove the existence of an optimal solution which is obtained by dynamic programming. As a result, we can formulate a series of problems dynamically consistent in the sense of [CCC$^+$11].

Inspiring again by [BET10], we study in §6.2 a dual problem for the value function we build through the extended Bellman equation for stochastic optimization problem with conditional expectation constraints.

6.1 From expectation constraints to extended Bellman equation

We are interested in the following problem:

$$\min_{(\mathbf{U},\mathbf{X})} \mathbb{E}\left[\sum_{t=0}^{T-1} L_t(\mathbf{X}_t, \mathbf{U}_t, \mathbf{W}_{t+1}) + K(\mathbf{X}_T)\right] \tag{6.1a}$$

where for all $t = 0 \dots T-1$, $\mathbf{X}_t$ (resp. $\mathbf{U}_t$) takes values in $\mathbb{X}_t$ (resp. $\mathbb{U}_t$) and $\mathbf{W}_{t+1} \in \mathbb{W}_{t+1}$, and $\mathbf{X}_T \in \mathbb{X}_T$, subject to dynamic constraints:

$$\mathbf{X}_0 = x_0,\ \mathbf{X}_{t+1} = g_{t+1}(\mathbf{X}_t, \mathbf{U}_t, \mathbf{W}_{t+1}), \tag{6.1b}$$

to measurability constraints:

$$\mathbf{U}_t \text{ is } \mathcal{F}_t\text{-measurable}, \tag{6.1c}$$

to an expectation constraint:

$$\forall i = 1 \dots d,\ \mathbb{E}\left[\gamma_i(\mathbf{X}_T)\right] \leq a_i. \tag{6.1d}$$

6.1.1 An equivalent problem with conditional expectation constraints

As it has been shown in the pioneer work of [BET10], and more specifically in [BEI10, Prop.5.1], and following [Gir10, Chap.5], Problem (6.1) is equivalent to the following problem:

$$\min_{(\mathbf{U},\mathbf{V},\mathbf{X},\mathbf{Z})} \mathbb{E}\left[\sum_{t=0}^{T-1} L_t(\mathbf{X}_t, \mathbf{U}_t, \mathbf{W}_{t+1}) + K(\mathbf{X}_T)\right] \tag{6.2a}$$

where for all $t = 0 \dots T-1$, $\mathbf{X}_t$ (resp. $\mathbf{U}_t$) takes values in $\mathbb{X}_t$ (resp. $\mathbb{U}_t$), $\mathbf{Z}_t = (\mathbf{Z}_t^i)_{i=1\dots d}$ and $\mathbf{V}_t$ belong to $\mathbb{R}^d$, and $\mathbf{W}_{t+1} \in \mathbb{W}_{t+1}$, and $\mathbf{X}_T \in \mathbb{X}_T$, subject to dynamic constraints:

$$\mathbf{X}_0 = x_0,\ \mathbf{X}_{t+1} = g_{t+1}(\mathbf{X}_t, \mathbf{U}_t, \mathbf{W}_{t+1}), \tag{6.2b}$$

$$\mathbf{Z}_0 = 0,\ \mathbf{Z}_{t+1} = \mathbf{Z}_t + \mathbf{V}_{t+1}, \tag{6.2c}$$

to measurability constraints:

$$\mathbf{U}_t \text{ is } \mathcal{F}_t\text{-measurable}, \tag{6.2d}$$

$$\mathbf{V}_t \text{ is } \mathcal{F}_t\text{-measurable}, \tag{6.2e}$$

to an almost sure final constraint:

$$\forall i = 1 \dots d,\ \gamma_i(\mathbf{X}_T) - \mathbf{Z}_T^i \leq a_i, \tag{6.2f}$$

and to the additional time constraints:

$$\mathbb{E}\left[\mathbf{V}_t | \mathcal{F}_t\right] \leq 0. \tag{6.2g}$$

6.1.2 Extended Bellman equation

We write the constraints (6.2f) through the characteristic function of the following set :

$$\mathit{\Gamma}_a := \Big\{(x,z) \in \mathbb{X}_T \times \mathbb{R}^d \,|\, \gamma(x) - z \leq a\Big\}$$

for some $a = (a_1, \dots, a_d) \in \mathbb{R}^d$.

Acccording to the dynamic programming principle stated in Equations (5.51), and due to the specific form of our problem, we define recursively (if it is possible) for $t = 0 \dots T$ the functions $V_t : \mathbb{X}_t \times \mathbb{R}^d \to \overline{\mathbb{R}}_+$ as following:

$$V_T^\gamma(x_T, z) = K(x_T) + \mathcal{X}_{\mathit{\Gamma}_a}(x_T, z) \tag{6.3a}$$

$$V_t^\gamma(x_t, z) = \min_{u_t \in \mathbb{U}_t} \min_{\mathbf{V} \prec \mathbf{W}_{t+1} : \mathbb{E}\big[\mathbf{V} | \mathcal{F}_t\big] \leq 0} \mathbb{E}\begin{bmatrix} L_t(x_t, u_t, \mathbf{W}_{t+1}) \\ +V_{t+1}^\gamma(g_{t+1}(x_t, u_t, \mathbf{W}_{t+1}), z + \mathbf{V}) \end{bmatrix} \tag{6.3b}$$

where $\mathcal{X}_{\mathit{\Gamma}_a}$ is the characteristic function (in the convex analysis sense) of $\mathit{\Gamma}_a$.

Therefore, the recursive step from $t+1$ to t is cut into two optimization problems.

Proposition 6.1.1 *Let us assume that for any t, $\mathbb{X}_t = \mathbb{X}_T$, and that for any $t = 0 \dots T-1$, $(x_t, u_t) \in \mathbb{X}_t \times \mathbb{U}_t \mapsto \mathbb{E}\left[L_t(x_t, u_t, \mathbf{W}_{t+1})\right] \in \mathbb{R}_+$ is an inf-compact function with respect to (x_t, u_t). Let us assume that $K : \mathbb{X}_T \to \mathbb{R}_+$ is an inf-compact function with respect to x_T. Let us assume that the following quantity $M := \min_{i=1\dots d} \inf_{x \in \mathbb{X}_T} g_i(x)$ is finite then the minimization Problem* (6.2) *either admits a solution using the recursive construction* (6.3) *or is not feasible.*

It proves at the same time that the dynamic programming (and the dynamic consistency) holds for the problem as soon as we consider the four processes $\mathbf{U}, \mathbf{X}, \mathbf{V}, \mathbf{Z}$, and that:

$$V_t(x_t, z) = \min_{(\mathbf{U}, \mathbf{X})} \mathbb{E}\left[\sum_{s=t}^{T-1} L_s(\mathbf{X}_s, \mathbf{U}_s, \mathbf{W}_{s+1}) + K(\mathbf{X}_T) \middle| \mathbf{X}_t = x\right] \tag{6.4a}$$

where for all $t = 0 \dots T-1$, $\mathbf{X}_t$ (resp. $\mathbf{U}_t$) takes values in $\mathbb{X}_t$ (resp. $\mathbb{U}_t$) and $\mathbf{W}_{t+1} \in \mathbb{W}_{t+1}$, and $\mathbf{X}_T \in \mathbb{X}_T$, subject to the same dynamic constraints (6.1b) and measurability constraints (6.1c):

$$\mathbf{X}_0 = x_0,\ \mathbf{X}_{t+1} = g_{t+1}(\mathbf{X}_t, \mathbf{U}_t, \mathbf{W}_{t+1}), \tag{6.4b}$$

to measurability constraints:

$$\mathbf{U}_t \text{ is } \mathcal{F}_t\text{-measurable}, \tag{6.4c}$$

to an expectation constraint:

$$\mathbb{E}\left[\gamma_i(\mathbf{X}_T) \middle| \mathbf{X}_t = x\right] \leq a + z. \tag{6.4d}$$

Proof. As in the Proof of Corollary 5.1.12, we introduce $\varphi : \mathbb{U}_{0:T-1} \times \mathbb{W}_{0:T} \to \mathbb{R}_+ \cup \{+\infty\}$, the total value of the cost when starting implicitely from x_0, we have applied the control u_t at time t, and where the sequence of noise processes $\mathbf{W}$ was equal to $w_{0:T}$.

For $t = T$, we set $\bar{\mathbb{U}}_T = (\mathbb{R} \cup \{+\infty\})^d$, and for $t = 1, \dots, T-1$, we set $\bar{\mathbb{U}}_t = \mathbb{U}_t \times (\mathbb{R} \cup \{+\infty\})^d$ and for $t = 0$, we set $\bar{\mathbb{U}}_0 = \mathbb{U}_0$.

We notice that by an inductive reasoning, we can prove that $V_t^\gamma(x_t, z) \geq V_t^0(x_t) + \mathcal{X}_M(z)$, where $\mathcal{X}_M$ is the characteristic function of the following set :

$$\left\{ z \in \mathbb{R}^d \mid \forall i = 1 \dots d,\ M + a_i \leq z_i \right\}, \tag{6.5}$$

and V_t^0 is the unconstrained value-function at time t. Note that this set is a compact set of $(\mathbb{R} \cup \{+\infty\})^d$.

We thus set:

$$\bar{\varphi} : (\bar{u}_{0:T}, w_{0:T}) \in \bar{\mathbb{U}}_{0:T} \times \mathbb{W}_{0:T} \mapsto \left(\begin{array}{c} \varphi(u_{0:T-1}, w_{0:T}) + \sum_{t=1}^{T} \mathcal{X}_M(\sum_{s=1}^{t} v_s) \\ + \mathbb{E}\left[\mathcal{X}_\gamma(\mathbf{X}_{0:T}^{u_{0:T-1}}, \sum_{t=1}^{T} v_t) \middle| \mathbf{W}_{0:T} = w_{0:T} \right] \end{array} \right) \in \mathbb{R} \cup \{+\infty\} \ . \tag{6.6}$$

As $\bar{\varphi}$ is inf-compact and the constraints are l.s.c, we can now apply Theorem 5.2.6 to get the result.

$\square$

6.2 Dual characterisation of extended Bellman equation

We want to compute in a quasi explicit way the minimization problem with the expectation constraint arising in Equation (6.3b).

6.2.1 A first abstract case

Let $(\mathbb{W}, \mathrm{Bor}(\mathbb{W}), \mathbb{P})$ be a non-atomic probability space, where $\mathbb{W}$ is a Lusin space. In this subsection d is an integer, and $\mathbb{U} = (\mathbb{R} \cup \{+\infty\})^d$. Let us define the functions $|.|_- : u \mapsto -u\mathbf{1}_{\{u \leq 0\}}$ and $\mathrm{Id} : u \mapsto u$, and let us define $\mathcal{R}_{\mathbb{P}}^-(\mathbb{W}, \mathbb{U})$ the subset of Young measures such that $I_{|.|_-}^{\mathbb{P}}$ is finite and $I_{\mathrm{Id}}^{\mathbb{P}}$ belongs to $\mathbb{R}_-^d$.

Assumption 6.2.1 *The function $\varphi : \mathbb{U} \times \mathbb{W} \to \mathbb{R} \cup \{+\infty\}$ is a $\mathbb{P}$ inf-compact function with respect to u such that if $I_\varphi^{\mathbb{P}}(\mu)$ is finite, then $I_{|.|_-}^{\mathbb{P}}(\mu)$ is finite.*

We are interested in the value of $a \in \mathbb{R} \cup \{+\infty\}$ where:

$$a := \inf_{\mu \in \mathcal{R}^-_{\mathbb{P}}(\mathbb{W},\mathbb{U})} I^{\mathbb{P}}_{\varphi}(\mu) \ . \tag{6.7}$$

Proposition 6.2.2 *Assume that Assumption 6.2.1 holds true. If $a = +\infty$, then the problem is not feasible. If a is finite, we have:*

$$a = \sup_{\lambda \in \mathbb{R}^d_+} \mathbb{E}\left[-\varphi^\star(-\lambda, \mathbf{W})\right] \tag{6.8}$$

where $\varphi^\star$ is any $\mathbb{P}$-version of:

$$\varphi^\star(\lambda', w) \in \mathbb{R}^d \times \mathbb{W} \mapsto \sup_{u \in \mathbb{R}} \lambda'.u - \varphi(u, w) \in \mathbb{R} \cup \{+\infty\} \ . \tag{6.9}$$

and . denotes the scalar product.

Proof. If $a = +\infty$, then by Corollary 5.3.2, it means that the minimization problem 6.7 is not feasible. Suppose that a is finite, we can restrict the infimum to the set of Young measures μ such that $I^{\mathbb{P}}_{\varphi}(\mu) \leq a+1$, which is a compact convex subset of $\mathcal{R}_{\mathbb{P}}(\mathbb{W}, \mathbb{U})$. We denote this set $\varphi^{\leq a+1}$. Note that due to the Assumption 6.2.1, for $\mu \in \varphi^{\leq a+1}$, then $I^{\mathbb{P}}_{|.|\_}(\mu)$ is finite, and thus $I^{\mathbb{P}}_{\mathrm{Id}}(\mu) = I^{\mathbb{P}}_{\mathrm{Id}+|.|\_}(\mu) - I^{\mathbb{P}}_{|.|\_}(\mu)$ is well defined as an element of $(\mathbb{R} \cup \{+\infty\})^d$, and we can consider only the constraint with $I^{\mathbb{P}}_{\mathrm{Id}}(\mu)$. By dualizing the constraint $I^{\mathbb{P}}_{\mathrm{Id}}(\mu) \in \mathbb{R}^d_-$, we have:

$$a = \inf_{\mu \in \varphi^{\leq a+1}} \sup_{\lambda \in \mathbb{R}^d_+} I^{\mathbb{P}}_{\varphi}(\mu) + \lambda.I^{\mathbb{P}}_{\mathrm{Id}}(\mu) \ . \tag{6.10}$$

We notice that $\varphi^{\leq a+1}$ is a convex compact set inside a metric space and that the following function:

$$f : (\mu, \lambda) \in \varphi^{\leq a+1} \times \mathbb{R}^d_+ \mapsto I^{\mathbb{P}}_{\varphi}(\mu) + \lambda.I^{\mathbb{P}}_{\mathrm{Id}}(\mu) \in \mathbb{R} \cup \{+\infty\} \ , \tag{6.11}$$

is convex and l.s.c partially in μ and concave in λ. Therefore, the assumptions of minimax Theorem of Ky Fan [Fan53, Th.2] are fulfilled. Consequently, we have:

$$\begin{aligned} a &= \sup_{\lambda \in \mathbb{R}^d_+} \inf_{\mu \in \varphi^{\leq a+1}} I^{\mathbb{P}}_{\varphi}(\mu) + \lambda.I^{\mathbb{P}}_{\mathrm{Id}}(\mu) \\ &= \sup_{\lambda \in \mathbb{R}^d_+} \inf_{\mu \in \varphi^{\leq a+1}} I^{\mathbb{P}}_{\varphi+\lambda.\mathrm{Id}}(\mu) \ , \end{aligned} \tag{6.12}$$

and since by Lemma 5.3.3, for any $\lambda \in \mathbb{R}^d_+$, $\varphi + \lambda.\mathrm{Id}$ is inf-compact with respect to u, we know by the results of Theorem 5.2.6 without expectation constraints that:

$$\inf_{\mu \in \varphi^{\leq a+1}} I^{\mathbb{P}}_{\varphi+\lambda.\mathrm{Id}}(\mu) = \mathbb{E}\left[-\varphi^\star(-\lambda, \mathbf{W})\right] = \int_{\mathbb{W}} -\varphi^\star(-\lambda, w)\mathbb{P}(dw) \ . \tag{6.13}$$

□

Corollary 6.2.3 *With the notations of Proposition 6.2.2. Let $p \in \mathbb{R}^d$, then we have :*

$$\inf_{\delta_{\mathbf{U}} \in \mathcal{R}^-_{\mathbb{P}}(\mathbb{W},\mathbb{U})} \mathbb{E}\left[\varphi(p+\mathbf{U},\mathbf{W})\right] = \sup_{\lambda \in \mathbb{R}^d_+} \mathbb{E}\left[-\lambda.p - \varphi^\star(-\lambda,\mathbf{W})\right] \ . \tag{6.14}$$

Proposition 6.2.4 *If it exists a function $q : \mathbb{W} \to \mathbb{R}^d$ which is $\mathbb{P}$-integrable such that φ is lower bounded by the characteristic function of the following set:*

$$\left\{(z,w) \in \mathbb{R}^d \times \mathbb{W} \mid \forall i = 1\dots d,\ z_i \geq q(w)\right\} \tag{6.15}$$

then Assumption 6.2.1 is satisfied, and the function depending on $p \in \mathbb{R}^d$ defined by Equation (6.14) *is convex and lower bounded by the characteristic function of the following set:*

$$\left\{z \in \mathbb{R}^d \mid \forall i = 1\dots d,\ z_i \geq \mathbb{E}\left[q(w)\right]\right\} \tag{6.16}$$

Proof. Let us denote $\mathcal{X}_q$ the characteristic function of the set defined by Equation (6.15). Let $\mu \in \mathcal{R}_{\mathbb{P}}(\mathbb{W},\mathbb{U})$ such that $\varphi(\mu)$ is finite, then $I^{\mathbb{P}}_{\mathcal{X}_q}(\mu) < \infty$, so it implies Assumption 6.2.1.

Let us notice that if for some $i \in \{1,\dots,d\}$, $z_i < \mathbb{E}\left[q_i(\mathbf{W})\right]$, then necessarily for any $\mu \in \mathcal{R}^-_{\mathbb{P}}(\mathbb{W},\mathbb{U})$, we have $I^{\mathbb{P}}_{\mathcal{X}_q}(\mu) = +\infty$. □

Proposition 6.2.5 *Assume that the problem depends parametrically on a variable ω which belongs to a Lusin space Ω. Then for any Borel probability measure $\mathbb{Q}$, it exists a control $\nu^\sharp : \Omega \mapsto \mathcal{R}_{\mathbb{P}}(\mathbb{W},\mathbb{U})$ such that if for a fixed ω the minimization problem* (6.7) *is finite, then $\nu^\sharp(\omega) \in \mathcal{R}^-_{\mathbb{P}}(\mathbb{W},\mathbb{U})$. Moreover, the Equation* (6.8) *holds $\mathbb{Q}$-almost surely.*

Proof. For the first statement, this is an application of Lemma 5.2.19. For the second statement, this is a combined application of Lemma 5.2.20 with Proposition 6.2.2. □

6.2.2 Applied case

We use the notations of §6.1.

Definition 6.1. *For any t, we define the p-convex conjugate of V_t as the function $V^\star_t : \mathbb{X}_t \times \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ such that:*

$$V^\star_t(x_t,\lambda) = \sup_{p \in \mathbb{R}^d} \lambda.p - V_t(x_t,p) \ . \tag{6.17}$$

This is a very standard notion of partial convex conjugate.

Proposition 6.2.6 *Under the Assumptions of 6.1.1, Equation* (6.3b) *can be replaced by:*

$$V_t(x_t,p) = \inf_{u_t\in\mathbb{U}_t} \sup_{\lambda\in\mathbb{R}^d_+} -\lambda.p + \mathbb{E}\left[L_t(x_t,u_t,\mathbf{W}_{t+1}) - V^\star_{t+1}\left(g_{t+1}\left(x_t,u_t,\mathbf{W}_{t+1}\right),\ -\lambda\right)\right]. \tag{6.18}$$

Proof. As we already mention, $V_t^\gamma(x_t,z) \geq V_t^0(x_t) + \mathcal{X}_M(z)$, where $\mathcal{X}_M$ is the characteristic function of the following set :

$$\left\{z\in\mathbb{R}^d \mid \forall i = 1\dots d,\ M + a_i \leq z_i\right\}, \tag{6.19}$$

and consequently, we are in the framework of Proposition 6.2.5. ◻

We now aim at looking at the equation satisfied by the p-convex conjuguate of V_t.

Proposition 6.2.7 *With the notations of Proposition 6.2.6, we get:*

$$V_t^\star(x_t,-\lambda) = \sup_{u_t\in\mathbb{U}_t} \mathbb{E}\left[-L_t(x_t,u_t,\mathbf{W}_{t+1}) + V^\star_{t+1}\left(g_{t+1}\left(x_t,u_t,\mathbf{W}_{t+1}\right),\ -\lambda\right)\right] \tag{6.20}$$

Proof. By Proposition 6.2.6, and by Definition 6.1, we get:

$$V_t^\star(x_t,-\lambda) = -\inf_{p\in\mathbb{R}^d} \inf_{u_t\in\mathbb{U}_t} \sup_{\lambda'\in\mathbb{R}^d_+} p.(\lambda-\lambda') + \mathbb{E}\left[L_t(x_t,u_t,\mathbf{W}_{t+1}) - V^\star_{t+1}\left(g_{t+1}\left(x_t,u_t,\mathbf{W}_{t+1}\right),\ -\lambda'\right)\right]\ . \tag{6.21}$$

We can intervert the infimum over p and the infimum over u_t since this is the infimum over a product space. The last point to prove that Equation (6.20) holds is the exchange of the supremum over λ' and the infimum over p. But, again, by considering that the infimum in p is taken over $(\overline{\mathbb{R}})^d$, Assumptions of [Fan53, Th.2] are fulfilled. A direct computation ends the proof. ◻

In conclusion, we have proved that the p-convex conjugate of V_t for a fixed λ is the value of a standard optimization problem. This remark was already mentionned in the introduction of [BET10].

Conclusion and further works

In this Chapter, we get some interesting results by dualizing the particular conditional expectation constraints of the Equation (6.3b), and we hope that they will lead to practical applications. Our work is very close to [BEI10], and we would like to make a clear connection between their results and our approach. We would like also to address in a future work, the question of numerical solutions for the minimization problems we have considered.

References

[AK05] L. Alili and A. E. Kyprianou. Some remarks on first passage of Lévy processes, the American put and pasting principles. *Ann. Appl. Probab.*, 15(3):2062–2080, 2005.

[Art01] Zvi Artstein. Compact convergence of σ-fields and relaxed conditional expectation. *Probability Theory and Related Fields*, 120:369–394, 2001.

[AS72] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions, With Formulas, Graphs, and Mathematical Tables,*. Dover Publications, Incorporated, 1972.

[Bal84] E.J. Balder. A general approach to lower semicontinuity and lower closure in optimal control theory. *SIAM journal on control and optimization*, 22(4):570–598, 1984.

[Bal89] E. Balder. On Prohorov's theorem for transition probabilities. *Sém. Anal. Convexe*, 19:Exp. No. 9, 11, 1989.

[Bal00] E. Balder. Lectures on Young measure theory and its applications in economics. *Rend. Istit. Mat. Univ. Trieste*, 31(suppl. 1):1–69, 2000. Workshop on Measure Theory and Real Analysis (Italian) (Grado, 1997).

[Bar04] K. Barty. *Contributions à la discretisation des contraintes de mesurabilité pour les problèmes d'optimisation stochastique.* PhD thesis, Ecole des Ponts ParisTech, 2004.

[BEI10] B. Bouchard, R. Elie, and C. Imbert. Optimal control under stochastic target constraints. *SIAM J. Control Optim.*, 48(5):3501–3531, 2009/10.

[Bel52] R. Bellman. On the theory of dynamic programming. *Proc. Nat. Acad. Sci. U. S. A.*, 38:716–719, 1952.

[Bel54] R. Bellman. The theory of dynamic programming. *Bull. Amer. Math. Soc.*, 60:503–515, 1954.

[Bel57] R. Bellman. *Dynamic programming.* Princeton University Press, Princeton, N. J., 1957.

[Ben84] A. Bensoussan. On the theory of option pricing. *Acta Applicandae Mathematicae*, 2(2):139–158, 1984. 10.1007/BF00046576.

[Ber96] J. Bertoin. *Lévy processes*, volume 121 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1996.

[Ber01] D. P. Bertsekas. *Dynamic programming and optimal control. Vol. I.* Athena Scientific, Belmont, MA, Second edition, 2001.

[BET10] B. Bouchard, R. Elie, and N. Touzi. Stochastic target problems with controlled loss. *SIAM J. Control Optim.*, 48(5):3123–3150, 2009/10.

[BL73] Henri Berliocchi and Jean-Michel Lasry. Intégrandes normales et mesures paramétrées en calcul des variations. In *Bulletin de la S.M.F.*, volume 101, pages 129–184. 1973.

[Bou71] N. Bourbaki. *Éléments de mathématique. Topologie générale. Chapitres 1 à 4*. Hermann, Paris, 1971.

[BR10] N. Bäuerle and U. Rieder. Markov decision processes. *Jahresber. Dtsch. Math.-Ver.*, 112(4):217–243, 2010.

[BS78] D. P. Bertsekas and S. E. Shreve. *Stochastic optimal control*, volume 139 of *Mathematics in Science and Engineering*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1978. The discrete time case.

[BS96] A. N. Borodin and P. Salminen. *Handbook of Brownian motion—facts and formulae*. Probability and its Applications. Birkhäuser Verlag, Basel, 1996.

[BX09] E. Bayraktar and H. Xing. Analysis of the optimal exercise boundary of american options for jump diffusions. *SIAM Journal on Mathematical Analysis*, 41(2):825–860, 2009.

[CC06] J. Chadam and X. Chen. A mathematical analysis of the optimal exercise boundary for american put options. *SIAM Journal on Mathematical Analysis*, 38:1613–1641, 2006.

[CCC+11] P. Carpentier, J.-P. Chancelier, G. Cohen, M. Lara, and P. Girardeau. Dynamic consistency for stochastic optimal control problems. *Annals of Operations Research*, pages 1–17, 2011.

[CCJZ08] X. Chen, J. Chadam, L. Jiang, and W. Zheng. Convexity of the exercise boundary of the american put option on a zero dividend asset. *Mathematical Finance*, 18:185–197, 2008.

[CJM92] P. Carr, R. Jarrow, and R. Myneni. Alternative characterizations of american put options. *Mathematical Finance*, 2(2):87–106, 1992.

[CRdFV04] C. Castaing, P. Raynaud de Fitte, and M. Valadier. *Young measures on topological spaces*, volume 571 of *Mathematics and its Applications*. Kluwer Academic Publishers, Dordrecht,

2004. With applications in control theory and probability theory.

[CRS71] Y. S. Chow, H. Robbins, and D. Siegmund. *Great expectations: the theory of optimal stopping*. Houghton Mifflin Co., Boston, Mass., 1971.

[CT04] R. Cont and P. Tankov. *Financial modelling with jump processes*. Chapman & Hall/CRC Financial Mathematics Series. Chapman & Hall/CRC, Boca Raton, FL, 2004.

[CV77] C. Castaing and M. Valadier. *Convex analysis and measurable multifunctions*. Lecture Notes in Mathematics, Vol. 580. Springer-Verlag, Berlin, 1977.

[DKW52] A. Dvoretzky, J. Kiefer, and J. Wolfowitz. The inventory problem. I. Case of known distributions of demand. *Econometrica*, 20:187–222, 1952.

[DM75] C. Dellacherie and P.A. Meyer. *Probabilités et potentiel*. Hermann, Paris, 1975. Chapitres I à IV, Édition entièrement refondue, Publications de l'Institut de Mathématique de l'Université de Strasbourg, No. XV, Actualités Scientifiques et Industrielles, No. 1372.

[Dyn63] E. B. Dynkin. Optimal choice of the stopping moment of a Markov process. *Dokl. Akad. Nauk SSSR*, 150:238–240, 1963.

[EK81] N. El Karoui. Les aspects probabilistes du contrôle stochastique. In *Ninth Saint Flour Probability Summer School—1979 (Saint Flour, 1979)*, volume 876 of *Lecture Notes in Math.*, pages 73–238. Springer, Berlin, 1981.

[EKPQ97] N. El Karoui, E. Pardoux, and M. C. Quenez. Reflected backward SDEs and American options. In *Numerical methods in finance*, volume 13 of *Publ. Newton Inst.*, pages 215–231. Cambridge Univ. Press, Cambridge, 1997.

[Eks04] E. Ekström. Convexity of the optimal stopping boundary for the american put option. *Journal of Mathematical Analysis and Applications*, 229(1):147–156, 2004.

[Evs76] I. V. Evstigneev. Measurable selection and dynamic programming. *Mathematics of Operations Research*, 1(3):pp. 267–272, 1976.

[Fan53] K. Fan. Minimax theorems. *Proceedings of the National Academy of Sciences of the United States of America*, 39(1):42, 1953.

[Fri88] A. Friedman. *Variational principles and free-boundary problems*. Robert E. Krieger Publishing Co. Inc., Malabar, FL, second edition, 1988.

[FS93] W. H. Fleming and H. M. Soner. *Controlled Markov processes and viscosity solutions*, volume 25 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, 1993.

[Gir10] P. Girardeau. *Résolution de grands problèmes en optimisation stochastique dynamique et synthèse de lois de commande*. PhD thesis, Université Paris-Est, 2010.

[GV11] O. E. Göttsche and M. H. Vellekoop. The early exercise premium for the American put under discrete dividends. *Math. Finance*, 21(2):335–354, 2011.

[Jac93] S.D. Jacka. Local , optimal stopping and semimartingales. *Annals of Probability*, 21:329–339, 1993.

[JJ12] M. Jeunesse and B. Jourdain. Regularity of the American Put option in the Black–Scholes model with general discrete dividends. *Stochastic Process. Appl.*, 122(9):3101–3125, 2012.

[JV11] B. Jourdain and M. Vellekoop. Regularity of the Exercise Boundary for American Put Options on Assets with Discrete Dividends. *SIAM Journal on Mathematical Finance*, 2:538–561, 2011.

[Kal02] O. Kallenberg. *Foundations of modern probability.* Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2002.

[KD00] H.J. Kushner and P.G. Dupuis. *Numerical methods for stochastic control problems in continuous time*, volume 24. Springer, 2000.

[Kim90] I.J. Kim. The analytical valuation of american options. *Review of Financial Studies*, 3(4):547–572, 1990.

[KRN65] K. Kuratowski and C. Ryll-Nardzewski. A general theorem on selectors. *Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys.*, 13:397–403, 1965.

[KS91] I. Karatzas and S. E. Shreve. *Methods of mathematical finance.* Springer, 1991.

[Lam95] D. Lamberton. Critical price for an American option near maturity. In *Seminar on Stochastic Analysis, Random Fields and Applications (Ascona, 1993)*, volume 36 of *Progr. Probab.*, pages 353–358. Birkhäuser, Basel, 1995.

[LL08] D. Lamberton and B. Lapeyre. *Introduction to stochastic calculus applied to finance.* Chapman & Hall/CRC Financial Mathematics Series. Chapman & Hall/CRC, Boca Raton, FL, second edition, 2008.

[LM08] D. Lamberton and M. Mikou. The critical price for the American put in an exponential Lévy model. *Finance and Stochastics*, 12(4):561–581, 2008.

[LM12a] D. Lamberton and M. Mikou. Exercise boundary of the american put near maturity in an exponential lévy model. *Finance and Stochastics*, pages 1–40, 2012.

[LM12b] D. Lamberton and M. Mikou. The smooth-fit property in an exponential Lévy model. *J. Appl. Probab.*, 49(1):137–149, 2012.

[LV03] D. Lamberton and S. Villeneuve. Critical price near maturity for an American option on a dividend-paying stock. *Ann. Appl. Probab.*, 13(2):800–815, 2003.

[McK65] H.P. McKean. Appendix: a free boundary problem for the heat equation arising from a problem of mathematical economics. *Ind. Management Rev.*, 6:32–39, 1965.

[Mey78] P.-A. Meyer. Convergence faible et compacité des temps d'arrêt d'après Baxter et Chacon. In *Séminaire de Probabilités, XII (Univ. Strasbourg, Strasbourg, 1976/1977)*, volume 649 of *Lecture Notes in Math.*, pages 411–423. Springer, Berlin, 1978.

[Mik58] V.S. Mikhalevich. Sequential bayes solutions and optimal methods of statistical acceptance control. *Theory Probab. Appl.*, 1:395–421, 1958.

[Mor02] E. Mordecki. Optimal stopping and perpetual options for Lévy processes. *Finance Stoch.*, 6(4):473–493, 2002.

[MS07] E. Mordecki and P. Salminen. Optimal stopping of Hunt and Lévy processes. *Stochastics*, 79(3-4):233–251, 2007.

[Myn92] R. Myneni. The pricing of the American option. *Ann. Appl. Probab.*, 2(1):1–23, 1992.

[Ped97] P. Pedregal. *Parametrized measures and variational principles*, volume 30. Birkhauser, 1997.

[Ped99] P. Pedregal. Optimization, relaxation and young measures. *BULLETIN-AMERICAN MATHEMATICAL SOCIETY*, 36:27–58, 1999.

[Pes05] G. Peskir. On the american option problem. *Mathematical Finance*, 15(1):169–181, 2005.

[Pha97] H. Pham. Optimal stopping, free boundary, and American option in a jump-diffusion model. *Appl. Math. Optim.*, 35(2):145–164, 1997.

[Pha09] H. Pham. *Continuous-time Stochastic Control and Optimization with Financial Applications*. Springer-Verlag Berlin Heidelberg, 2009.

[PP11] T. Pennanen and A. Perkkiö. Stochastic programming without duality gaps. Technical report, Quelque part, 2011.

[Pra90] L. Pratelli. Sur le lemme de mesurabilité de Doob. In *Séminaire de Probabilités, XXIV, 1988/89*, volume 1426 of *Lecture Notes in Math.*, pages 46–51. Springer, Berlin, 1990.

[PS06] G. Peskir and A.N. Shiryaev. *Optimal stopping and free-boundary problems*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2006.

[Put94] M.L. Puterman. *Markov decision processes: Discrete stochastic dynamic programming*. John Wiley & Sons, Inc., 1994.

[Rie78] U. Rieder. Measurable selection theorems for optimization problems. *Manuscripta Math.*, 24(1):115–131, 1978.

[Roc76] R. T. Rockafellar. Integral functionals, normal integrands and measurable selections. In *Nonlinear operators and the calculus of variations (Summer School, Univ. Libre Bruxelles, Brussels, 1975)*. Springer, Berlin, 1976.

[RW98] R. T. Rockafellar and R. J.-B. Wets. *Variational analysis*, volume 317 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1998.

[RY91] D. Revuz and M. Yor. *Continuous martingales and brownian motion*. Springer, 1991.

[Shi08] A. N. Shiryaev. *Optimal stopping rules*, volume 8 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2008. Translated from the 1976 Russian second edition by A. B. Aries, Reprint of the 1978 translation.

[Sne52] J. L. Snell. Applications of martingale system theorems. *Trans. Amer. Math. Soc.*, 73:293–312, 1952.

[ST02a] H. M. Soner and N. Touzi. Dynamic programming for stochastic target problems and geometric flows. *J. Eur. Math. Soc. (JEMS)*, 4(3):201–236, 2002.

[ST02b] H. M. Soner and N. Touzi. Stochastic target problems, dynamic programming, and viscosity solutions. *SIAM J. Control Optim.*, 41(2):404–424, 2002.

[Thi81] L. Thibault. Espérances conditionnelles d'intégrandes semi-continus. *Ann. Inst. H. Poincaré Sect. B (N.S.)*, 17(4):337–350, 1981.

[Vil] C. Villani. Integration et analyse de fourier. http://math.univ-lyon1.fr/homes-www/villani/Cours/iaf-2007.html.

[VM76] P. Van Moerbeke. On optimal stopping and free boundary problems. *Arch. Ration. Mech. Anal.*, 60:101–148, 1976.

[VN06] M.H. Vellekoop and J.W. Nieuwenhuis. Efficient pricing of derivatives on assets with discrete dividends. *Applied Mathematical Finance*, 13(3):265–284, September 2006.

[VN11] M. H. Vellekoop and J. W. Nieuwenhuis. An integral equation for American put options on assets with general dividend processes. *Stochastics*, 83(4-6):555–567, 2011.

[Wal47] A. Wald. *Sequential Analysis*. John Wiley & Sons Inc., New York, 1947.

[WW49] A. Wald and J. Wolfowitz. Bayes solutions of sequential decision problems. *Proc. Nat. Acad. Sci. U. S. A.*, 35:99–102, 1949.

[You37] LC Young. Generalized curves and the existence of an attained absolute minimum in the calculus of variations. *Comptes Rendus de la Société des Sciences et des Lettres de Varsovie, classe III*, 30:212–234, 1937.